

NCBI_TaxID=65699;
 RN SEQUENCE FROM N.A.
 RC STRAIN-22491 / SEROGROUP A / SEROTYPE 4A;
 RX MEDLINE-20222556; PubMed-10761919;
 RA Parkhill J., Achtman M., James K.D., Bentley S.D., Churcher C.,
 RA Klee S.R., Morrell G., Basham D., Brown D., Chillingworth T.,
 RA Davies R.M., Davis P., Devlin K., Feltwell T., Hamlin N., Holtroyd S.,
 RA Jagers K., Leather S., Moule S., Mungall K., Quail M.A.,
 RA Rajandream M.A., Rutherford K.M., Simmonds M., Skelton J.,
 RA Whitehead S., Spratt B.G., Barrall B.G.;
 RT "Complete DNA sequence of a serogroup A strain of *Neisseria meningitidis* 22491.";
 RL Nature 404:502-506(2000).
 DR EMBL: AL162753; CAB83949.1;
 KW Transferase; Acyltransferase; Complete proteome.
 SQ SEQUENCE 298 AA; 33510 MW; 2A09946670E283C2 CRC64;

alignment_scores:
 Quality: 1580.00 Length: 298
 Ratio: 5.302 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
 US-09-303-518D-569 x Q9JYV7 ..

Align seg 1/1 to: Q9JYV7 from: 1 to: 298

```

1 ATGTTGCTTACAAATTCAGGCTGTTCCCTTTGGCAACGCCATGCA 50
1 MetPheArgLeuGlnPheArgLeuPheProPheArgTrpAlaMetHis 17
51 CATCTGTTGAGCGCGCTGCTCAATGCTCTCCCTGCTGGCGCTTCT 100
17 stlleuLeuThrAlaLeuLeuLysCysLeuSerLeuLeuSerC 34
101 GTCTGCACACGCTGGGAAACCGGCTCGGACATCTGGCGTTTACCTTTA 150
34 yLeuHisThrLeuGlnAsnArgLeuGlnHisLeuAlaPheThrLeu 50
151 AAGGAAGACCGCGCGCATCTGCTCCCAATATGCTCAGGACGGATGAA 200
51 LysGlnAspArgAlaArgIleValAlaAsnMetArgGlnAlaGlyMetAs 67
201 TCCCGACCCCAAAACGCTCAAGCCGTTTTCGGAACGCGCAAAAGCG 250
67 nProAspProLysThrValLysAlaValPheAlaGluThrAlaLysGly 84
251 GTTTGGAACCTGCCCGCGGTTTTCAGAAAACGGAAGACATAGAACA 300
84 yLeuGluLeuAlaProAlaPhePheArgLysProGlnAspIleGluThr 100
301 AGTTCAAAAGCGGTACAGCGCTGGGACATGTCAGCAGCGCTTGGACA 350
101 MetPheLysAlaValHisGlyTrpGlnHisValGlnGlnAlaLeuAsp 117
351 ACAGAAAGGCGTGCATTATCATCAGCGCGCATCGGACGCTAGATTGG 400
117 shsGlnGlyLeuLeuPheIleThrProHisIleGlySerTrpAspLeu 134
401 GCGGACGCTACATCAGCGACAGCTTCCGCTCCGCTGACCGCATGTC 450
134 yGlyArgTrpLysIleSerGlnGlnLeuProPheProLeuThrAlaMet 150
451 AAACCGCCCAAAATCAACGATAGACAAATCATCGCAGCGCGGAGGT 500
151 LysProProLysIleLysAlaIleAspLysIleMetGlnAlaGlyArg 167
501 TCGCGGCAAGGAAACCGCGCTACGACATACGACGAGGTCAAAACA 550
167 lArgGlyLysGlyLysThrAlaProThrSerIleGlnGlyValLysGln 184

```

```

551 TCATCAAGCCCTCGGTCGGGCGAAGCAACCAATGCTCCGCCGACCAC 600
184 lLleLysAlaLeuArgSerGlyGlnAlaThrIleValLeuProAspHis 200
601 GTCCCTCCCTCCCAAGAGCGGGGAGGCGGTATGGGTGATTTCTTGG 650
201 ValProSerProGlnGlnLysGlyGlnGlyValTrpValAspPhePheG 217
651 CAACCTGCTTATACATGACGCTGGCGCAAAATTTGGCAGCAGTCAAG 700
217 yLysProAlaTrpThrMetThrLeuAlaAlaLysLeuAlaHisValLys 234
701 GCGTGAACCCCTGTTTCTGCTGCGAAGCGCTGCTGGCGGACAGGT 750
234 yValLysThrLeuPhePhePheCysGlyArgLeuProGlyLysGln 250
751 TTGGATTGACATCGCGCGCGCTCCCAAGGGAATTGAACGGGACAAAG 800
251 PheAspLeuHisIleArgProValGlnGlyLysLeuAsnGlyAspLysAl 267
801 CCATGATGCGCGCGCTGTCACCGCAATGCCGAATTTGATACGCCGTT 850
267 ahIsAspAlaIleValPheAsnArgAsnAlaGluTrpTrpIleArgArg 284
851 TTCGACGCGAGTATCTGTTATGTACACCGCTACAAATGCCG 894
284 heProThrGlnTrpLeuPheMetLysAsnArgTrpLysMetPro 298

```

seq_name: sp_bacteriap:Q9JY07

```

seq_documentation_block:
ID Q9JY07 PRELIMINARY; PRT; 298 AA.
AC Q9JY07;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE HTRB/MSBB FAMILY PROTEIN.
GN NMB1801.
OS Neisseria meningitidis (serogroup B).
OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
OX NCBI_TaxID=491;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-MC58 / SEROGROUP B;
RX MEDLINE-20175755; PubMed-10710307;
RA Tettelein H., Saunders N.J., Heidelberg J., Jeffries A.C., Nelson K.E.,
RA Eisen J.A., Ketchum K.A., Hood D.W., Peden J.F., Dodson R.J.,
RA Nelson W.C., Gwin M.L., DeBoy R., Peterson J.D., Hickey E.K.,
RA Haft D.H., Salzberg S.L., White O., Fleischmann R.D., Dougherty B.A.,
RA Mason T., Ciecko A., Parksey D.S., Blair E., Clifton H., Vamathevan J.,
RA Cotton M.D., Utterback T.R., Khouri H., Qin H., Vamathevan J.,
RA Gill J., Scariato V., Maignant V., Pizze M., Grandi G., Sun L.,
RA Smith H.O., Fraser C.M., Moxon E.R., Rappuoli R., Venter J.C.;
RT "Complete genome sequence of Neisseria meningitidis serogroup B strain MC58."
RL Science 287:1809-1815(2000).
DR EMBL; AE002530; AAF42138.1; -.
KW Complete proteome.
SQ SEQUENCE 298 AA; 33492 MW; 013437BC4DD6508F CRC64;

```

alignment_scores:
 Quality: 1577.00 Length: 298
 Ratio: 5.292 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 99.664

alignment_block:

US-09-303-518D-569 x Q9JY07 ..

Align seg 1/1 to: Q9JY07 from: 1 to: 298

```

1 ATGTTGCTTACAAATTCAGGCTGTTCCCTTTGGCAACGCCATGCA 50

```

```

|||||
1 MetPheArgLeuGlnPheArgLeuPheProLeuArgThrAlaMetH1 17
51 CATCGCTGTAACGGCCGCTCAAAATGCGTCCCTGCTGCGCTTCT 100
|||||
17 sIleuLeuThrAlaLeuLeuLysCysLeuSerLeuLeuProLeuSer 34
101 GTCTGCACAGCGCTGGAAACCGGCTGACATCGGCTTTTACCTTTA 150
|||||
34 yLeuHnIsthrLeuGlyAsnArgLeuGlnHsLeuAlaPheTyrLeu 50
151 AAGGAAGACCGCGCGCATGCTGCCAATATGCGTCAGGACGATGAA 200
|||||
51 LysGlnAspArgAlaArgIleValAlaAsnMetArgGlnAlaGlyLeu 67
201 TCCGACCGCCAAACGGCTCAAAACGGCTTTGGCGAAACGGCAAGCG 250
|||||
67 nProAspProLysThrValLysAlaValPheAlaGlnThrAlaLysGly 84
251 GTTTGGAACCTGCCCCCGCTTTTTCAGAAACCGGAAACATAGAAAC 300
|||||
84 yLeuGlnLeuAlaProAlaPhePheArgLysProGlnAspIleGlnThr 100
301 ATGTTCAAAGCGGTACACGGCTGGGAACATGTCAGAGGCTTTGACAA 350
|||||
101 MetPheLysAlaValHnIsthrGlnHsValGlnGlnAlaLeuAspLy 117
351 ACAGGAAGGCTGCTATTCATCACCGCGCATGCGGACGATGATTTGG 400
|||||
117 sHsIsthrGlyLeuLeuPheHnIsthrProHnIsthrGlySerTyrAspLeu 134
401 GCGGACCGCTACATCACGACGACGCTTCCGCTCCGCTGACCGCATGTAC 450
|||||
134 yGlyArgTyrIleSerGlnGlnLeuPheProLeuThrAlaMetTyr 150
451 AAACCGCGGAAATCAAGCGATAGACAAATATGTCAGCGGCGAGGT 500
|||||
151 LysProLysIleLysAlaIleAspLysIleMetGlnAlaGlyArg 167
501 TCCGCGGAAAGGAAACCGCGCTTACAGATACAGAGGCTCAACAA 550
|||||
167 LArgGlyLysGlyLysThrAlaProThrSerIleGlnGlyValLysGln 184
551 TCATCAAGCGCTGCTGCGGGAAGCAACATGCTGCTGCGCGCACAC 600
|||||
184 LelLeuAlaLeuArgSerGlyGlnAlaThrIleValLeuProAspHn 200
601 GTCCCGCTCCCTCAGAAAGCGGGAAGCGCTATGATGATTTCTTGG 650
|||||
201 ValProSerProGlnGlnGlyGlyGlnGlyValItrPvalAspPheG 217
651 CAACCGCTTACCATGACGCTGCGGCAAAATGCGACACGTAAG 700
|||||
217 yLysProAlaTyrThrMetThrLeuAlaLysLeuAlaHnIsthrVal 234
701 GCGTGAACCGCTTTTCTGCTGCAAGCGCTGCTGCGGCAACAGT 750
|||||
234 yValLysThrLeuPhePheCysCysGlnAlaGlyLeuProGlyGly 250
751 TTGCATTTGCATCCGCGCGCTCCAGGGGAATTAAGAGCGGCAAGC 800
|||||
251 PheAspLeuHnIsthrArgProValGlnGlyLeuAsnGlyAspLysAl 267
801 CCATGATGCGCGCTGTTCAACCGCATGCGGAAATGATTTGGATACG 850
|||||
267 nHnIsthrAlaAlaValPheAsnArgAsnAlaGlnTyrTrrPleArgArg 284
851 TTCCGACGAGTATCTGTTATGTACAACCGCTCAAAATGCGC 894
|||||
284 nProThrGlnLysLeuPheMetTyrAsnArgTyrLysMetPro 298

```

seq_name: sp_bacteria:Q9EYMO

```

seq_documentation_block:
ID Q9EYMO PRELIMINARY; PRT; 337 AA.
AC Q9EYMO;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE PUTATIVE HNRB PROTEIN.
OS Acinetobacter sp. BD413.
OC Bacteria; Proteobacteria; gamma subdivision; Moraxellaceae;
OX Acinetobacter.
NCBI_TaxID=104611;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BD413;
RX MEDLINE=21318707; PubMed=11425734;
RA Friedrich A., Hartsch T., Averhoff B.;
RT "Natural Transformation in Mesophilic and Thermophilic Bacteria:
RT Identification and Characterization of Novel, Closely Related
RT Competence Genes in Acinetobacter sp. Strain BD413 and Thermus
RT thermophilus HB27."
RL Appl. Environ. Microbiol. 67:3140-3148(2001).
DR EMBL; AF320001; AAG34711.1; -.
DR InterPro; IPR000515; BPD.transp.
DR PROSITE; PS00402; BPD_TRANSF_INN_MEMBER; UNKNOWN_1.
SQ SEQUENCE 337 AA, 38795 MW; 6E998AFEDICEB140 CRC64;

```

alignment_scores:

Quality:	Ratio:	Length:
293.00	2.170	193
Percent Similarity: 69.948		Gaps: 5
		Percent Identity: 35.233

alignment_block:

US-09-303-518D-569 x Q9EYMO

Align seg 1/1 to: Q9EYMO from: 1 to: 337

```

313 GTACAGCGGCGGGAACATGTCAGACGCTTTGGCAAAACGAAAGCGT 362
|||||
100 ValGlnGlyAlaAspPheHnIsthrAlaIleAlaGlnAsnLysGlyI 116
363 GCTATTCATCACCGCGCACATGCGGACGATGATTTGGCGGACGCTCA 412
|||||
116 eValLeuIleValProHnIsthrPheGlyThrTrpLysIleMetAsnAlaTrpC 133
413 TCAGCGACGACGCTCCGCTCCCGCTGACCGCATGTACAAACCGCGAAA 462
|||||
133 ysaIaGlnTyrThr....AspMetThrIleLeuTyrLysProValLys 147
463 ATCAAGCGATAGACAAATATCATGACGGGCGGAGGCTTGGCGGCAAG 512
|||||
148 AspLysAspAlaAspArgPheValArgGlnAlaAlaArgSerArgGlnAl 164
513 AAAACCGCGCGCTACACGATACAGGGGCTCAAAACATCATCAAGCCC 562
|||||
164 aLysLeuValProThrAspGlnSerGlyValArgGlnIlePheLysAla 181
563 TGCCTTCGCGGCAACCAACATGCTGCTGCGGCAACGATGCTCCCT 612
|||||
181 euLysGlnIleGlyThrValIleLeuProAspHn....ThrPro 195
613 CAAGAAGCGGGAAGGCGTATGCGGATTTCTTGGCAACGCTGCTA 662
|||||
196 AsnValGlyGlyGlu.....MetValAspTyrPheGlyIleProLeuAl 210
663 TACCATGACGCTGCGGCGCAAAATTTGGCACACGTCGTAAGGCTGAACCC 712
|||||
210 aSerSerAsnLeuSerAlaLysLeuIsthrGlnLysThrLysAlaLysAla 227
713 TGTTTTCTGCTGCGAAGCGCTGCGGCAAGGCTTTCGATTGGAC 762
|||||
227 euPheLeuTyrAlaIleArg...AsnGlnAsnHnIsthrPheAspMetPhe 242

```


seq_documentation_block:

ID 091785 PRELIMINARY; PRT: 295 AA.
AC 091785:
DT 01-MAR-2001 (Tremblrel. 16, Created)
DT 01-MAR-2001 (Tremblrel. 16, last sequence update)
DT 01-OCT-2001 (Tremblrel. 18, last annotation update)
DE PROBABLE 2-OH-LAUROYLTRANSFERASE.
GN PA0011.
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas
OX NCBI_TaxID=287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 15692 / PA01;
RC MEDLINE=20437337; PubMed=10984043;
RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warriner P.,
RA Hickey M.J., Brinkman F.S.L., Hutnagle W.O., Kowalik D.J., Lagrou M.,
RA Garber R.L., Collier L., Tolentino E., Westbrook-Wedman S., Yuan Y.,
RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;
RT *Complete genome sequence of Pseudomonas aeruginosa PA01, an
RT opportunistic pathogen.";
RL Nature 406:959-964(2000).
DR EMBL: AE004441; AAG03401.1;
KW Transferase; Complete proteome.
SQ SEQUENCE 295 AA; 33001 MW; D6D1710EC6AD42FF CRC64;

alignment_scores:

Quality: 259.00 Length: 286
Ratio: 1.463 Gaps: 7
Percent Similarity: 61.888 Percent Identity: 25.874

alignment_block:

US-09-303-518D-569 x 091785

Align seg 1/1 to: 091785 from: 1 to: 295

```
55 CTTGTGACCGCCCTGCTCAATGCTCTGCTGCGCTTTCGTCT 104
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
8 LeuValValGlyAlaLeuArgLeuPheAlaLeuLeuProTyrPargAlaVal 24
105 GCACACGCTGGGAAACCGCTCGACATCGCGCTTTTAACTTAAAG 154
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
24 GlnGlyValGlyAlaGlyIleGlyTyrPleuMetTyrPlysLeu.....P 39
155 AAGACCGCGCGCGCATCGTCCCAATATGCGTCAGCAGCATGATCC 204
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
39 rGlnArgSerArgGlnValValArgIleAsnLeuSerLysCysPhePro 55
205 GAC.....CCCAAAAGCGTCAAGCCGTTTGGCGGA 236
|||||:|||||:|||||:|||||:|||||:|||||:
56 GluLeuSerGlnThrGlnLeuGlnLeuValGlyGlnSerLeuMetAs 72
237 AACGGCAAAAGCGGTTTGAAGTTCGCCCGGTTTTCAGAAACCGG 286
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
72 PileGlyArgThrLeuThrGlnSerAlaCysAlaTyrPileTyrProProG 89
287 AAGACATAGAAACAATGTCAAAGCGGTACAGCGCTGGCAACATGTCAG 336
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
89 LysSerLeuArgTyrIleArgGlnValGlnIleGlyMetGlnValLeuGln 105
337 CAGCTTTGGACAAACAGGAGGCTGTATTCATCAGCCCGGCATCGG 386
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
106 GlnAlaLeuAlaSerGlyAspGlyLeuValGlyIleThrSerHisLeuGln 122
387 CAGCTAGAT...TTGGGCGGAGCTACATACGACGAGCGTTCGCTTC 433
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
122 YAsnTyrGlnValLeuAsnHisPheTyrCysSerTyrAlaLysProIle. 138
```

seq_name: sp_bacteriap:09HYZ8

seq_documentation_block:

ID 09HYZ8 PRELIMINARY; PRT: 312 AA.
AC 09HYZ8:
DT 01-MAR-2001 (Tremblrel. 16, Created)
DT 01-MAR-2001 (Tremblrel. 16, last sequence update)
DT 01-OCT-2001 (Tremblrel. 18, last annotation update)
DE PROBABLE LAUROYL ACYLTRANSFERASE.
GN PA3242.
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas
OX NCBI_TaxID=287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 15692 / PA01;
RC MEDLINE=20437337; PubMed=10984043;
RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warriner P.,
RA Hickey M.J., Brinkman F.S.L., Hutnagle W.O., Kowalik D.J., Lagrou M.,
RA Garber R.L., Collier L., Tolentino E., Westbrook-Wedman S., Yuan Y.,
RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;
RT *Complete genome sequence of Pseudomonas aeruginosa PA01, an
RT opportunistic pathogen.";
RL Nature 406:959-964(2000).
DR EMBL: AE004747; AAG06630.1;
KW Transferase; Acyltransferase; Complete proteome.
SQ SEQUENCE 312 AA; 36063 MW; 0BDE771D75F3C186 CRC64;

alignment_scores:

```
434 CGCTGACCGCCATGTACAAACCGCCGAATTCAAAGCATAGCAAAATC 483
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
139 .....IlePheTyrArgProTyrPheLysAlaValAspGlnLeu 152
484 ATGCAGCGCGGAGGTTTCGCGGCAAGAAACCGCGCTTCGCGCAT 533
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
153 LeuLysLysGlnArgValGlnLeuGlyAsnArgValAlaProSerThrPyr 169
534 ACAAGGGGTCAACAATTCATCAAGCCCTGCTGGCGGCAAGCAACCA 583
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
169 GlnGlyIleLeuSerValIleLysGlnValLysLysGlyGlyCysValG 186
584 TCGTCTGCGCCGACCAAGCGTCCCTCCCTCAAGAGCGGGAGCGCTA 633
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
186 LysLeuProAlaAsp.....ProGlnProAlaArgThrAla...GlyLeu 199
634 TGGGTGATTTCTTTCGCAACCTGCTATACCATGACGCTGGCGCAAA 683
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
200 PheValProTyrLeuGlnIleThrAlaLeuIleSerLysPheValProGln 216
684 ATTGGCACAGTCAAGCGGTGAACCCCTGTTTCTCTGCTCGCAACGCC 733
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
216 nLeuLeuSerArgGlyLysAlaArgGlyAlaPhePheHisAlaValArgL 233
734 TGGCTGCGGACAGGTTTCGATTTGACATCCGCCCGCTCCAAAGGGA 783
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
233 euProAspGlySerGlyTyrLysValIleLeuGlnAlaAlaProAlaAsp 249
784 TTG...AACGGACAAAGCCCATGATGCGCGCGCTTCAACCGCAATGC 830
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
250 MetTyrAspLysAspLeuGlnValSerValAlaAlaMetSerArgGlnLe 266
831 CGAATTTGATACGCGCTTTCGACGACAGTATCTGTTTGTATGTAAC 880
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
266 uAlaLysTyrValArgAlaTyrProSerGlnTyrMetTyrPheMetLys 283
881 GCTACCAA 888
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
283 rgPheLys 285
```

[REDACTED]

```
seq_documentation_block:
ID 09JYV2 PRELIMINARY; PRT; 289 AA.
AC 09JYV2;
DT 01-OCT-2000 (TrEMBLrel. 15 Created)
DT 01-OCT-2000 (TrEMBLrel. 15 Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 15, Last annotation update)
DE HTB/MSB FAMILY PROTEIN.
ID NM61418.
OS Neisseria meningitidis (serogroup B).
OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria
NCBI_Taxid=491;
[1]
```

RP SEQUENCE FROM N.A.
RC STRAIN-MC58 / SPROGROUP B;
RX MEDLINE-20175755; PubMed-10710307;
RA Tedellin H., Saunders N.J., Heidelberg J., Jeffries A.C., Nelson K.E.
RA Nelson J.A., Ketchum K.A., Hood D.W., Feden J.F., Dodson J.D., Hickey E.J.
RA Nelson W.C., Gwinn M.L., Deboy R., Peterson J.D., Hickey R.J.,
RA Haft D.H., Salzberg S.L., White O., Fleischmann R.D., Dougherty B.A.,
RA Mason T., Clecko A., Parksey D.S., Blair E.C., Clifton H., Clark E.B.,
RA Cotton M.D., Uitterlidge T.R., Khouri H., Qin H., Vamathevan J.,
RA Gill J.J., Scariato V., Massignani V., Pizzo M., Grandi G., Sun L.,
RA Smith H.O., Fraser C.M., Moxon E.R., Rappunli R., Venter J.C.,
RT "Complete genome sequence of *Neisseria meningitidis* serogroup B strain
RL MC58.";
RL Science 287:1809-1815(2000).
DR EMBL; AF002491; AAF1779.1; -.
DR TIGR; NM01418; -.
KW Complete proteome.
SQ SEQUENCE 289 AA: 33847 RT: 1

alignment_scores:

Quality:	204.50	Length:	286
Ratio:	1.225	Gaps:	11
Percent Similarity:	58.392	Percent Identity:	26.923

Alignment_block:
US-09-303-518D-569 x Q9JYV2

Align seg 1/1 to: Q9JYV2 from: 1 to: 289

[illegible]

```

Align seg 1/L co. cctcccccgcggcggcggttttcctgacacagcgtgg 116
67 ctgcctcaaatgccctccatccggcgccggcgttctctgacacagcgtgg 116
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
7 valleutyValleuinlneuleuprophealaleuinhlsyleuL 23
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
117 AAACCGGCTGGAGCAATCTGGCGTTTAAAGGAAGACCGCGGC 166
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
23 aaapleuthrglyeuLeuLeuValtyrleueValtyrprotagargargi 40
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
167 GCATCTGCCCAATATCCGTACAGCAGCATTCGCCGCCCAACAAG 216
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
40 leclgluileashleualalsycyspheProglutrpasbgilylsy 56
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
217 GTCAAAGCCGTT.....TTTGCGGAAACGGCAAAAGCGGTT 254
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
57 ArglysthnValleuLysglInhisheylshisMetalaLysleumetle 73
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
255 GGAACCTGGCCCCGCGTTCGAAAACCGGAAGACTATGAACAATGT 304
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
73 ugiutyrglyleuTytrTYrAlarProlaClglyargleuYsserleuv 90
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
305 TCAAAGCGTACACGGCTGGAAACATGTGCACAGCGTTTGGCAACAAC 354
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
90 alary...TYrAghsnlYshIStyLeuaspRAspAlaleuLaalaglY 105
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
355 GAAGGCGCTGATTTCATCACGCGGCAACATGGCAGCTAGATTGGGCG 404
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
106 GlulyVallellelleuTytrPronhsphethrlarPheglumetaIva 122
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
405 ACCTACATCGACGACAGCTTCGGTCCGCGTACCAGCCCATATAAAC 454
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
122 lYrAlaleuasnGlinaSPval.....ProleuIlseImetyrSerH 137

```

RC STRAIN=20365717. PubMed=10910347. Agencio M.,
RX Simpson A.J.G., Reimach F.C., Arruda P., Abreu F.A., Baptista C.S.,
RA Alvares A.L., Alves L.M.C., Araya J.E., Bala G.S., Briones M.R.S.,
RA Alvarenga R., Bonaccorsi E.D., Bordin S., Bove J.M., Carraro D.M.,
RA Barros M.H., Bonaccorsi A.A., Camargo L.E.A., Carraro D.M., Carrier H.,
RA Bueno M.R.P., Camargo A.A., Costa F.F., Costa M.C.R., Costa-Neto C.M.,
RA Colantuoni N.B., Colombo C., Costa F.F., Docena C., El-Dorriy H.,
RA Colantuoni L.P., Cristofani M., Dias-Neto E., Docena C., El-Dorriy H.,
RA Coutinho A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,
RA Facchini L.P., Ferreira A.J.S., Franco M.C., Frohme M., Furian L.R.,
RA Fraga J.S., Franca S.C., Franco M.H.S., Gomes S.L., Gruber A.,
RA Garner M., Goldman G.H., Goldman M.H.S., Kempster E.L., Kitzajima J.P.,
RA Ho P.L., Hohensei J.D., Junqueira M.L., Lambais M.R., Leite L.C.C.,
RA Klieger J.E., Kurnaa E.E., Laigret F., Lambais M.R., Machado J.A.,
RA Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Martino C.Y.,
RA Machado M.A., Madella A.M.B.N., Medeira H.M.F., Matsukuma A.Y.,
RA Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma C.B.,
RA Menck C.F.M., Miracca E.C., Miyaki C.Y., Montefiro-Vitorello C.B.,


```

55 uclyagtrhglalaglultrhlanislausprhrllephesrtyra 72
      ::::: ||| ||| |||
281 AACGGAGACATAGAACATGTTTCAAGCGGTACAGCGCGGGAACAT 330
      ||||| ::::: |||||
72 sptroblusnleualatrhglnarqmetgluvalargluleuugln 88
      ||||| ::::: |||||
331 GTGCACGACGCTTGTGCAACACGAAAGGCTGTATTCATCCGCGCA 380
      ||| ::::: |||
89 phephalalalearghsprsglgnlprogluleuilepheseralani 105
      ||| ::::: |||
381 CATGGCAGCATGATTTGGCGGAGCGGTACATCAGCCAGACGTCCT 430
      ||||| ::::: |||||
105 sleualasnttrpgluleuoproala...llecsalaglnulysphegyl 121
      ||||| ::::: |||||
431 TCCGCGTACCGCATGTACAAACCGCGAAATCAAGCGATAGACAA 480
      ::::: ||||| ::::: |||||
121 euglulhtrhtrhalavalpheargprrprobsnlnproalaalaglaln 137
      ||| ::::: |||
481 ATCATGCAG.....GCGGCGAGGTTGCGGCAAGG 512
      ::::: ||| ::::: |||
138 leuvalglngluvalargarglsthmetglgylleualalasergl 154
      ||| ::::: |||
513 AAAAACCGCGCGGTACAGCATACAGGCGTCAACAAATCATCAAGCCG 562
      ||| ::::: |||||
154 yproglalavalphealmetglnlgluvalglu..... 166
      ||| ::::: |||
563 TGCCTTGGCGGAGCAACCATGCTGCTGCGGACGACGTCCTCCCT 612
      ||||| ::::: |||||
167 ..Argglglnlslleuglgluleuileasprlnlspherhtr..... 180
      ||||| ::::: |||||
613 CAAGAGCGCGGAGAGCGGTATGGGTGATTTCTGCGCAACGTCGCTA 662
      ::::: ||||| ||| |||||
181 .....Arggluvalavalthrpherhegluargprovalle 193
      ||||| ::::: |||||
663 TACCATGACGCTGCGGCAAAATTCGCACAC.....GTCA 697
      ::::: ||| ::::: |||
193 uvalasnproleuenglulysleualatrhlanislasprysprovalh 210
      ||||| ::::: |||||
698 AAGCGGTAAACCGCTGTTTCTGCTGCGAAGCGCTGCGTGGGCAAA 747
      ||||| ::::: |||||
210 lsglualatargvalval.....Arglsglglgyl 220
      ||||| ::::: |||||
748 GATTTCGATTTCACATCCGCC..... 771
      ||| ||||| ::::: |||
221 ArgphleuenglulthrProproleuhsprleuProatrgluproas 237
      ||||| ::::: |||||
772 .....GTCAAGGGGATTCAGCGGACAAAGCCCATGATG 808
      ||||| ::::: |||||
237 polyleuilegluvalglnglualameglu..... 247
      ||||| ::::: |||||
809 CCGCGGTGTCACCGCAATGCGCAATATGATAGCGGTTTCCGACG 858
      ::::: ||| ||| |||||
248 ....Alamethlthargvalillegluglultrpvalarglulhnsprogly 262
      ||||| ::::: |||||
859 CAGTATCTGTTATGTACACCGGTACAA 888
      ||||| ::::: |||||
263 glntprleuMetlnlsarglgtprarg 272
      ||||| ::::: |||||
seq_name: sp_bacteria:Q9L7U9
seq_documentation_block:
ID Q9L7U9 PRELIMINARY; PRT; 318 AA.
AC Q9L7U9;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-OCT-2000 (Tremblrel. 15, Last annotation update)
DE Lipid A ACYLTRANSFERASE.
GN MSBB.
OS Haemophilus ducreyi.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Haemophilus.
OX NCBI_TaxID=730;

```

```

RN [1]
RP SEQUENCE FROM N.A.
RA Munson R.S. Jr., Bozue J.A.
RT "Identification of the msbB gene in Haemophilus ducreyi."
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF219260; AAF33777.1;
KW Transferase; Acyltransferase.
SO SEQUENCE 318 AA; 36799 MW; EAC1F82FB88BE37A CRC64;

alignment_scores:
  Quality: 172.50      Length: 262
  Ratio: 1.113        Gaps: 10
  Percent Similarity: 59.160      Percent Identity: 24.427

alignment_block:
US-09-303-518D-569 x Q9L7U9

Align seg 1/1 to: Q9L7U9 from: 1 to: 318

132 TCTGCGCTTTTACCTTTTAAAGAGACCGCGCGCATCGTCCGCA.. 179
      ||||| ||| ||||| ::::: |||
56 Serasnlytleuylslgylslglnlthrtmetleuglilegl 72
      ||||| ::::: |||||
180 .....TATGCGTACGCGCATGATGATCCGACCCCAAA 213
      ::::: ||||| ::::: |||||
72 easnleucystyrcyspneprogluttrpseralalsprysarglaln. 88
      ||||| ::::: |||||
214 ACGGTCAAGCGCTTTTTCGGAACGCGCAAAAGCGGTTTGAACCTGC 263
      ::::: ||| ::::: |||
89 valilleglulmetrpherlelnsvalalaglthrmetleuglilegl 105
      ||||| ::::: |||||
264 CCGCGGCTTTTCAAGAAACCGGAACATAGAAATGTTCAAGCGG 313
      ||||| ::::: |||||
105 yclullealavalargservalalalnlsleuclnlsargsergluphei 122
      ::::: ||||| ::::: |||||
314 TACAGCGGTGGAGACATGTCAGACGCTTGGCAAAACAGAGGCTG 363
      ::::: ||| ::::: |||
122 le...glyleuasnlytleuylslglnlthrtmetleugllylslsle 137
      ||||| ::::: |||||
364 CTATTCATCAGCGCGCACATCGCGACGTACGATTTGGCGGCGGTACAT 413
      ::::: ||||| ||| |||||
138 lleuileuvalprohlsrthrtrpsertleaspralasergl...llel 153
      ||||| ::::: |||||
414 CAGCGACGACGCTCCGCTCCGCTGACCGCATGTACAAACCGCGCAAA 463
      ||||| ::::: |||||
464 TCAAGCGATAGACAAATCATGACAGCGGCGCATGTCGCGGCAAAAG 513
      ::::: ||| ::::: |||
170 snlaleuvalasprtrleuTrpasnalatrharglulargrphleglly 186
      ||||| ::::: |||||
514 AAAACCGCGCGCTACCGACATACAGGCGGTCAACAAATCATCAAGCCCT 563
      ||||| ::::: |||||
187 Lys...Methlsrthrarglinsngllyllelslsppropheuasnlsprl 202
      ||||| ::::: |||||
564 GCGTTCGCGGAGCAACCATGCTGCTGCCCGGACCGACGTCCTCCCTC 613
      ||||| ::::: |||||
202 earglsgllyasmetcyltytrpherleuproasr..... 6214
      ||||| ::::: |||||
614 AAGAGCGGCGGAA.....GCGGTATGGGTGATTTCTTCGCAAACT 657
      ::::: ||||| ::::: |||||
214 lnsasprheglglugluleuaservaltytralasnprhephalatrngln 230
      ||||| ::::: |||||
658 GCGTATCATGACGCTGCGGCAAAATTCGACACGTCGCAAGCGGTGA 707
      ||||| ::::: |||||
231 lysalatrheuprogluleuasnlysmetalargvalalalsnagl 247
      ||||| ::::: |||||
708 AACCGTG...TTTTCTGCTGCGAACCGCTGCTGCGGCAAGGTTTCG 754
      ::::: ||| ::::: |||
247 uvalilleuowetpetherproilletyrasnalaglulysglval...tyrg 263
      ||||| ::::: |||||
755 ATTGCACATCGCGCGCGTCAAGGGAATTCGAAGCGCGCAAGCCCAT 804
      ||||| ::::: |||||

```

```

263 1mmeGluileuProAlaile...GluPheSerGlySerSerAlaGln 278
805 GATGCCCGCGTGTCAACCGCAATGCCGATATTGGATACCGCGTTTCC 854
279 SerAlaArgGluMetAsnLysValIleGluIrrPheValHisLysAsnPr 295
855 GACGCAATGCTGTATTATGTACAAACCGCTACAA 888
295 oAlaGlnIrrValIrrPheLeuArGleuArLys 306
seq_name: sp_bacteriap:084013

```

```

seq_documentation_block:
ID 084013 PRELIMINARY; PRT; 455 AA.
AC 084013;
DT 01-NOV-1998 (TREMBLrel. 08, Created)
DT 01-NOV-1998 (TREMBLrel. 08, last sequence update)
DE ACYLTRANSFERASE.
GN HTRB OR CT010.
OS Chlamydia trachomatis.
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=813;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=D/OW-3/CX;
RX MEDLINE=99000809; PubMed=9784136;
RA Stephens R.S., Kaiman S., Lammel C.J., Fan J., Marathe R., Aravind L.,
RA Mitchell W.P., Olinger L., Tatusov R.L., Zhao Q., Koonin E.V.,
RA Davis R.W.;
RT "Genome sequence of an obligate intracellular pathogen of humans:
RT Chlamydia trachomatis."
RL Science 282:754-759(1998).
DR EMBL; AE001275; AAC67600.1;
KW Transferase; Acyltransferase; Complete proteome.
SQ SEQUENCE 455 AA; 52058 MW; 0404B6326C67ACCF CRC64;

```

```

alignment_scores:
Quality: 172.50 Length: 342
Ratio: 1.033 Gaps: 12
Percent Similarity: 48.830 Percent Identity: 23.977

alignment_block:
US-09-303-518d-569 x 084013 ..
Align seg 1/1 to: 084013 from: 1 to: 455

```

```

22 CTGTTTCCCTTTCGAACCGCATGCAATCCGTTGAGCGCCG.. 69
||||| ||||| ||||| ||||| ||||| |||||
2 LeuPheLysArGLeuArGThrGlyLysLleLeuValAspHisLeuVa 18
70 .....CTCAATGCCCTCTCCCTGCTGCGCTT 97
18 IYrGlyLeuGlyLeuGlyValLeuThrLleLeuArGLeuLeuProArG 35
98 CCGTCTGCACACGCTGGGAACCGGCTCGACATCTGGCGTTTACTT 147
|| ||||| ||||| ||||| ||||| ||||| |||||
35 eSerLeuArGLeuPheSerLysGlyLleGlyThrAlaLeuPheTyPhe 51
148 TTAAGAGAGACCGCGCGCATGCTGCGCAATG..... 183
||||| ||||| ||||| ||||| ||||| |||||
52 IleSerAspValArGlyThrAlaLeuThrAsnLeuAlaLeuAlaPhePr 68
184 .....GTCACGCA..... 192
68 oGluLysSerPheAlaGluArGlyGlnIleAlaArGHisSerValGln 85
193 .....GCGATGAATCCCGACCCCAAGAGGTCGCAAA 222
85 ImeLleIleThrPheValGluLeuAlaThrValAspLysPheAlaLys 101

```

```

223 GCGGTTTTTGGCGAAACGGCAAAAGCGGTTTGAACCTTGGCC...GC 269
||| ||| ||| ||| ||| ||| |||
102 HisLysAspGluMetLleAlaIleAlaThrSerGluAspAlaProIuL 118
270 GTTTTTCAGAAACCGGACACATAGAACATGTTCAACCGGTACAG 319
||||| ||||| ||||| ||||| ||||| |||||
118 YPhePhe.....ProGlnIuValSerSerGlnGlnIuLeu..... 130
320 GCTGGACAAATGTGCACAGCTTTTGGACAAACACAGCGCGCTATTC 369
||||| ||||| ||||| ||||| ||||| |||||
131 .....AspHisPhePheSerArGLeuAspArgGlnIuLAlaIleLeu 145
370 ATCACGCCGCACATCGCGACGCTACGATTTGGCGGACGCTACAGCA 419
||| ||| ||||| ||||| ||||| |||||
146 PheCysGlyHisGlnAlaAsnTrGlnLeuProPheLeuTyLleThy 162
420 GCACCTTCCTCCGCTGACCGGCAATGTCACAAACCGCGCAAAATCAAG 469
||| ||| ||||| ||||| ||||| |||||
162 sArG.....TyrrProGlyLeuAlaPheAlaLysProValLysAsnArG 177
470 CGATGACAAATCATGACAGCGCGGCGGTTTCGCGCAAGAAACAC 519
||||| ||||| ||||| ||||| ||||| |||||
177 rGleuAsnGlnLysLleIleSerLeuArGLeuSerPheGlnIuLysLle 193
520 GCGGCTACACGATACAGAGGCTCAACAAATCATCAAGCCCTGCTTC 569
||| ||| ||||| ||||| ||||| |||||
194 ValrProGln...AsnAlaLysGlnAlaLeuArGAlaLeuHisAr 209
570 GGGCGAAGCAACCATGCTCCGCGGACGAGCTCCGCTCCGCTCAAGAG 619
||||| ||||| ||||| ||||| ||||| |||||
209 gGlyGluValAlaGlyLleValGlyAspGlnValLeuLeuSerSerGln 226
620 GCGGGAGAGCGGTATGGGTGATTTCTTCGCAACCTGCTATACATG 669
||| ||| ||||| ||||| ||||| |||||
226 ySer.....TyrrProLeuPheGlySerGlnAlaPheThr 238
670 ACGCTGGCGCAAAATTTGGCACACGTCAAAGCGCTGAACCCGTGTTT 719
||| ||| ||||| ||||| ||||| |||||
239 ThrSerProAlaLeuAlaIrrLysThrLysLysThrValIleAlaVa 255
720 CTGCTGCGAAGCGCTGCTGGCGGCAAGGTTTGATTTGCACATCCGCC 769
||| ||| ||||| ||||| ||||| |||||
255 lAlaLleTyrrArGlyLysProAsnGlyAsnTyrrLeuVal..... 268
770 CCGTCCAAAGGGAATTGAACGGCGCAACGCCATGATCCGCGCTGTC 819
||| ||| ||||| ||||| ||||| |||||
269 .....ProSerLysAlaPhe 273
820 AACCGCAATGCCGATATTGGATACG..... 846
||| ||| ||||| ||||| ||||| |||||
274 HisAlaAsnThrGlnLeuSerLleArGLeuSerThrGlnGlnLeuMetAs 290
847 .....CGTTT..... 862
290 rArGLeuMetArGrPheLeuGlnLysGlyLleThrCysLysProGlnIuLnt 307
863 ATCTGTTTATACAAACCGCTACAA 888
||||| ||||| ||||| ||||| ||||| |||||
307 rPleuTrPheuHisLysArGTrPrLys 315
seq_name: sp_bacteriap:092CCL1

```

```

seq_documentation_block:
ID 092CCL1 PRELIMINARY; PRT; 290 AA.
AC 092CCL1;
DT 01-MAY-1999 (TREMBLrel. 10, Created)
DT 01-MAY-1999 (TREMBLrel. 10, last sequence update)
DE 01-DEC-2001 (TREMBLrel. 19, last annotation update)
DE LIPID A BIOSYNTHESIS LAUROYL ACYLTRANSFERASE (HTRB).
GN RP718.
OS Rickettsia prowazekii.
OC Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
OC Rickettsiaceae; Rickettsiinae; Rickettsia.

```



```

227 leProcySglnlleIleargThrlysglySerlyrPhelysValIleVal 243
766 CGCCCG...GTCCAGGGGAATTGACGGCGACAAAGCCATGATGCC.. 810
244 HisProIleuIuysPhegluInThrGlyAspAsnlyValasPcsty 260
811 .....GCCGCTTTCACCGCAATGCCAATATTGATGATGCGCGT 850
260 AsnIleMetPheThrIleAsnGlnIleuGlyGlnIleValysGlnA 277
851 TTCGAGACGATATCTGTTATGTACACCGCTACAAA 888
277 snProIaGlnIleThrPheThrPheHisAsnArgTrlyps 289
seq_name: sp_bacteriap:Q9PL30

```

```

seq_documentation_block:
ID   Q9PL30      PRELIMINARY;   PRT;   453 AA.
AC   Q9PL30;
DT   01-OCT-2000 (TREMBlrel. 15, Created)
DT   01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DE   01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE   LIPID A BIOSYNTHESIS LAUROYL ACYLTRANSFERASE, PUTATIVE.
GN   TC0278
OS   Chlamydia muridarum.
OC   Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
OX   NCBI_TaxID=83360;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MOBN / NIGG;
RX   MEDLINE=20150255; PubMed=10684935;
RA   Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
RA   White O., Hickey E.K., Peterson J., Ullerbach T., Berry K., Baas S.,
RA   Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,
RA   Givhn M., Nelson W., DeBoy R., Kolonay J., McClarty G., Salzberg S.L.,
RA   Eisen J., Fraser C.M.;
RT   "Genome sequences of Chlamydia trachomatis MOBN and Chlamydia
RT   pneumoniae AR39.";
RL   Nucleic Acids Res. 28:1397-1406(2000).
DR   EMBL; AE002295; AAF39146.1; -.
DR   TIGR; TC0278; -.
KW   Complete proteome.
SQ   SEQUENCE 453 AA; 51463 MW; 6221515A00A093FF CRC64;

```

```

alignment_scores:
Quality: 164.50      Length: 336
Ratio: 0.962        Gaps: 14
Percent Similarity: 50.893      Percent Identity: 24.702

```

alignment_block:

US-09-303-518D-569 x Q9PL30 ..

Align seg 1/1 to: Q9PL30 from: 1 to: 453

```

22 CTGTTTCCCTTTCGAAACGGCATGACATCCCTGTTG..... 60
||||| |||||.....
2 LeuPheIysMetLeuArgSerGlyIlyValIleuValAspRHisLeuVal 18
61 .....ACCGCTGCTGCATTAATGGCTCTCCCTGCTGCGCTTT 97
||||| |||||.....
18 IlyGlyLeuGlyIleuGlyIleLeuThrIleLeuArgLeuLeuProArgS 35
98 CCGTCTGCACAGCGTGGGAAACCGCGTCGACATGCGCTTTTACCTT 147
|| |||||.....
35 eSerLeuGlnLeuPheGlyIlySgIyLeuGlyThrIlePheIyVal 51
148 TTAAGAAGAACCGCGCGCATGCTGCGCAATATGCGTCAGGAGGAT 197
||||| |||||.....
52 IleSerAspPheArgIyThrAlaLeuThrAsnLeuAlaLeuAla..... 66
198 GAATCCGACACCC.....AAAAAGC 217

```

```

67 PheProGlyIuysSerPheThrGluArgTyrgInIleAlaLeuIysSerV 83
||||| |||||.....
218 TCMAAGCCCTT.....TTGGCGAAAGCGCAAA..... 246
||||| |||||.....
83 alGlnGlnValIleIleThrPheIleGluLeuAlaThrValasPlySph 99
247 .....GGCGTTTGCACTTGCCCG 266
||||| |||||.....
100 AlaIysHisIleAspGlnIleIleThrIleAlaIleSerSerGluAspAlaPr 116
267 C...GGCTTTTCACAAAACCGAAGACATAGAAACAATGTTCAAGCGG 313
||||| |||||.....
116 ocIuGlyPhePhe.....ProGluGlyValSerSerGlnGlnIleuLeu. 130
314 TCACGCGCTGGGACATGTGACAGCGCTTGACAAACAGAGGCGTG 363
||||| |||||.....
131 .....AsnAsnPhePheSerArgLeuAspGlnGlnGlyAla 143
364 CTATTCATCAGCGCCGACATCGGACGTAGCATTTGGCGGACGCTACAT 413
||||| |||||.....
144 IleuPheCysGlyHisGlnAlaAsnTrpIleuPheProPheLeuTyrl 160
414 CAGCAGCAGCGCTCCGCTCCCGCTGACCGCGCATGTACAAACCGCGAAA 463
||||| |||||.....
160 eThrIysArg.....TyrrProGlyLeuAlaPheAlaIysProValIlysa 175
464 TCMAAGCATATGACAAATATCATGACAGCGGCGAGGCTTCGCGCAAGA 513
||||| |||||.....
175 snProArgLeuAsnArgIysIleIleSerLeuArgGlnUserPheGlnGly 191
514 AAACCGCGCTTACCGATACAGGCGTCAAAACATCATCAAGCCCT 563
||||| |||||.....
192 LysIleValProProGln...AsnAlaIleAsnGlnAlaLeuAlaGlyAla 207
564 GCGTTCGGGCGCAACGATCTCTCTCCCGACGACGCTCCCTCCCTG 613
||||| |||||.....
207 uHisIysGlyIleuValIyAlaGlyIleValIyAspGlnValLeuIysers 224
614 AAGAAGCGGGGGAAGCGGTATGGGTGATTTCTTGGCAAACTGCTAT 663
||||| |||||.....
224 eArgIlyrSer.....TyrrProLeuPheGlySerGlnAlaPhe 236
664 ASCATGACGCTGGCGCAAAATGGCACAGGTCAAAAGCGTGAACCCCT 713
||||| |||||.....
237 ThrThrIrrSerProAlaLeuLeuAlaIyThrIyThrArgIyProValI 253
714 GTTTTCTGCTGCGAAGCGCTGCTGGCGAACA..... 747
||||| |||||.....
253 eAlaValAlaIleTyrrArgGlnProAsnGlyAsnTyLeuValIyValPro 270
748 .....GGTTGATTTGCAATCCGCGCCCGCCGCAAGG 780
||||| |||||.....
270 eIlySAlaPheTyrrAlaAsnThrGlyIleuSerIleArgGlnSerThrIgl 286
781 GAATTGACGCGCGCAAAAGCCATGATGCCGCGGTTCACCGCAATGC 830
||||| |||||.....
287 GlnLeuMet...AspIys.....LeuMetArgPheLe 296
831 CGAATATGATAGCGCGCTTTCCGACGCGATGTCTGTTATGATACAC 880
||||| |||||.....
296 uGlyIySgIyIleAlaCysIySProGlnIleThrIleuIyThrIleuIy 313
881 GCTACAAA 888
||||| |||||.....
313 rGTrIyIys 315
seq_name: sp_bacteriap:Q9CME5
seq_documentation_block:
ID   Q9CME5      PRELIMINARY;   PRT;   313 AA.
AC   Q9CME5;
DT   01-JUN-2001 (TREMBlrel. 17, Created)

```



```

129 ACATCGGCGTTTACCTTTAAGAGAGACGGCGCGTGTGCCA 178
    |||.....|||
58 LysileValalalargylsalaarghlsalalysile 75
179 ATATGCGTCAGCGCAGCATGAAT.....CCGACCCCAAG 216
    |||.....|||
75 snleuglnTYCysPheProHisTrpGlnlaGlnarglnIntlr 91
217 GTCAAGCGGTTTTCGGAAGAGGCGGTTTGGACTTCCCC 266
    |||.....|||
92 llelueglnmetrhevalilevalinvalmetleugllel 108
267 CGCCTTTTCAAGAAACCGGAGACATAGAACATGTTCAGCGGTAC 316
    |||.....|||
108 uilealileargserlysrghisleuglnarg...SerCysPheT 124
317 AGCGCTGGAAACATGTGACAGCGTTTGGACAAACAGAGCGCTGA 366
    |||.....|||
124 hrglueglnhislehisarglalargluInglYlysanielle 140
367 TTATCAAGCGCGCATCGGACCTAGATTGGCGGACGCTACATAG 416
    |||.....|||
141 leuuevalProhisalatrpalaleasplasercly...llelle 156
417 CGAGAGCTTCGCTCCGCTGACCGGCATGTACAAACCGCGAAATCA 466
    |||.....|||
156 uhlsthrhisglymerPheProMetThrSerMetTYrAsnProHisArgasp 173
467 AAGCGATAGCAAAATCATGACGAGCGGCGGATTCGCGGCAAGAA 516
    |||.....|||
173 roleuvalasprleuprleuprhrhealalarglIntrPheglYlys 189
517 ACCGCGCTACAGCATACAGGGTCAAAACATCATCAAGCGCTCG 566
    |||.....|||
190 MethisalarglIn...AsnglYleuYsProPheleuAsnhsVally 205
567 TTGCGGCGAGCAACCATGCTCTCGCGCGACGCTCCCTCCCTCAAG 616
    |||.....|||
205 sglInglYasprmetglYTYrTYrleuprAspIn.....AspT 218
617 AAGCGCGGGA...GCGCTATGGTGATTTCTTCGCAACCTGCTCAT 663
    |||.....|||
218 YrclYAlaGlnleuSerValPheValasprPhealathrTYrlysla 234
664 ACCATGACCTGCGGCAAAATGTGCACAGCTC.....AAGCGCTGA 707
    |||.....|||
235 ThrleuproglylleasnlYsillealargleualalysalaalaval 251
708 AACCTGTTTTCTGCTGGAACGCTCTGCGGACAGGTTTCGAT 757
    |||.....|||
251 eProMetPheProArgTYrAsnAlaleSerCly.....LysTYrGln 266
758 TGCACATCGCGCGCTCCAGAGGGAATGAAGCGGACAAAGCCATGAT 807
    |||.....|||
266 leGlnllehisProalmet...ThrleuSerAsprAsprProglIngl 281
808 GCCGCGCTGTCAACCGCAATGCGCAATATGTAGCGCTTTCCGAC 857
    |||.....|||
282 AlaargalalmetasnglnIngllelSerPhevalThrAlatrProal 298
858 GCAGATCTGTTATGTACAAACGCTACAAA 888
    |||.....|||
298 aGlnTYrValTrpIleuargTYrleuYs 308

```

seq_name: sp_bacteria:Q9L7Q3

seq_documentation_block:

```

ID Q9L7Q3 PRELIMINARY: PRT: 317 AA.
AC Q9L7Q3;
DT 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TReMBLrel. 15, Last annotation update)
DE HTRB.

```

```

GN HTRB.
OS Haemophilus ducreyi.
OC Bacteria: Proteobacteria: gamma subdivision: Pasteurellaceae;
OC Haemophilus
OX NCBL:Taxid=730;
RN [1]
RP SEQUENCE FROM N.A.
RA Munson R.S. Jr., Bozue J.A.;
RT "Identification of the htrb gene in Haemophilus ducreyi.";
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF221008; AAF34642.1;
SQ SEQUENCE 317 AA; 36521 MW; 8359FE1D5945612B CRC64;

```

alignment_scores:

```

Quality: 161.50 Length: 286
Ratio: 1.003 Gaps: 12
Percent Similarity: 56.294 Percent Identity: 25.874

```

alignment_block:

US-09-303-518d-569 x Q9L7Q3 ..

Align seg 1/1 to: Q9L7Q3 from: 1 to: 317

```

64 GCCCTGCTCAATGCTCTCCCTGCGCGCTTCTGTCACAGCT 113
    |||.....|||
32 AlalePheleYsleuileuCYsleuProTYrProileleuVallysl 48
114 GGGAAACCGGCTCGACATCTGCGCTTTTACCTTTAAGAGAGACGG 163
    |||.....|||
48 eglYargserleuAlalyseu...PheSerThrleuSerPheglYlys 64
164 CGGCGATC...GTGCGCAATATGCTCAAGCAGCGCATCAATCCGAC 207
    |||.....|||
64 rgarYlealevalalalarYargasnlleuAlaleuCYsPheProAspr 80
208 ...CCCAAAACGTCAAACCGCTTTTGGGAAACGGCAAGCGGT. 252
    |||.....|||
81 serAlaGlnGlnlleGlnAlaleleuileuYlsasnlValGlnserVal 97
253 .....TTGAACTTGCCCGCGCTTTTTCAGAAACCGGAGACA 292
    |||.....|||
97 ymetAlalelleGlnThrGlnYmetAlatrPheTrpSerAsprlysrY 114
293 TAGAAACATGTTCAAACGCGTACAGCGCTGGAACATGTGCACAGCT 342
    |||.....|||
114 le...leuYsTrpserAsnlleGlnYleuGlnHisleuGln... 128
343 TTGACAAACAGAGGCGCTGATTCATCAGCGGCGACATCGGACGTA 392
    |||.....|||
129 ...ProlysglyThrGlylleuPheValGlyAlaHisPheleuThrle 144
393 CGATTTGGCGGCGGCTACATCAGCGGACGCTTCCGCTCGACG 442
    |||.....|||
144 uGlnleuGlyAlaArgyleValGly...leuHisleGlnGlylleG 159
443 CCATGTACAAACCGCGCAATTCAAAGCGATAGCAACATGACGAGCG 492
    |||.....|||
159 lyValTYrArgProAsnAsprAsnProleuYsIntPrilleGlnPheYs 175
493 GCGAGGCTGCGCGCAAGGAAACCGCGCTACAGCATACAGAGGCT 542
    |||.....|||
176 glYargyleArg...SerAsnlyAlaMetleuAsprTYrlysrAspr 190
543 CAACAAATCATCAAGCGCTGCTGCGGCAAGGACATGCTGCTCG 592
    |||.....|||
190 uargGlymetlleargAlaleuAsprlysglyGlnThrPleTrpTYrAla 207
593 CGGACAGCGTCCCTCCCTCAAGAGGCGGGAAGCGATGAGGCGAT 642
    |||.....|||
207 roAsprHisAsprTYrGlyGlnIn...AsnserValPhevalPro 220
643 TTCTTGGCAACCTGCTTACCATGACGCGTGGCGCAAAATGTGCACA 692

```

```

|||||:
221 PhePheAlaValGluGlnAlaCysThrThrAlaGlySerHisMetLeu 237
693 CGTCAAGGCGTGAACCCCTGTTT.....TTCTGCTCGAAGCGCTGC 736
237 uArgSerAlaProAsnSerIleValAlaProPheThrProIleArgHisT 254
737 CTGGCGGACAGAGTTTCGATTTCACATCCGCCCTCCAGGGCAATTG 786
254 yAspPheSerGlyTyrThrValLysIleSerProAlaVal..AspHe 269
787 AACGCG.....GACAAAGCCATGATGCCCGCTTCACCCGCAATGC 830
270 AsnAspGlyGlnAsnLysIleGluThrAlaArgLeuMetAsnLysVal 286
831 CGAATATGATACGCCGTTTCCGAGCGAGTATCTGTTATGTACACG 880
286 lGluThrGluIleuLysAlaGlnThrGlnTyrMetThrPheHisArg 303
881 GCTACAAA 888
303 rghPheLys 305

```

seq_name: sp_bacteria:09AKF0

```

seq_documentation_block:
ID 09AKF0 PRELIMINARY; PRT; 290 AA.
AC 09AKF0;
DT 01-JUN-2001 (TREMblrel. 17, Created)
DT 01-JUN-2001 (TREMblrel. 17, Last sequence update)
DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
DE HTRB PROTEIN.
GN HTRB.
OS Rickettsia rickettsii.
OC Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
OC Rickettsiaceae; Rickettsiae; Rickettsia.
OX NCBI_Taxid=783;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=84-21C;
RX MEDLINE=21219194; PubMed=11319266;
RA Andersson J.O., Andersson S.G.E.;
RT "Pseudogenes, junk DNA and the dynamics of Rickettsia genomes.";
RL Mol. Biol. Evol. 18:829-839(2001).
DR EMBL: A1793329; CAC3715.1;
SQ SEQUENCE 290 AA; 33550 MW; 9FC5D35E5CBE89A CRC64;

```

alignment_scores:
 Quality: 154.00 Length: 201
 Ratio: 1.252 Gaps: 8
 Percent Similarity: 61.194 Percent Identity: 23.881

alignment_block:
 US-09-303-518d-569 x 09AKF0 ..

Align seg 1/1 to: 09AKF0 from: 1 to: 290

```

310 GGGGTACAGCGGTGGAGACATGACGAGGCTTGGACAAACAGAGG 359
|||||:|||||:
101 AAlaIleIleGlyIleGlnAsnIleLys.....LysLeuGlnG 113
360 G.....CTGTATTCATCACCGCCGACATCGGACGATGATGGGCG 403
|.....|.....|.....|.....|.....|.....|.....|.....|
113 yGlnProPheLeuLeuPheSerGlyHisPheAlaAsnTrpAspIleSerI 130
404 GACGCTACATCAGCCAGAGCTTCGCTCCGCTGACCGCCATGTACAA 453
||:|||||:
130 lLysIleLeuHisLysPheTyrPro..LysValAlaValIleTyrArg 145
454 CGCCGCAAAATCAAGCATAGCAAAATCATGACGCGCGGAGGTTGCG 503
||:|||||:
146 LysAlaAsnAsnProTyrValAlaAsnLysLeuValAsnGluSerArgAla 162

```

```

504 CGGCAAGGAAAAACCCGCTTACCACATACAGAGGGGTCAACAAATCA 553
||:|||||:
162 yAspLysLeuArgLeuIleProLysGlyProGlnGlySerArgAlaLeu 179
554 TCMAAGCCCTCGTTTCGGGACAGCAACATCTCTGCCGACCCAGCTC 603
179 AlArgAlaIleuLysGluSerGlnAlaIleValaMetLeuValAsp..... 193
604 CCGTCCCTCAAGAGCGCGGAGAGCGGTATGGTGCATTTCTTCGGCA 653
194 .....GlnLysMetAsnAspGlyIleGluValaProPheLeuGlyH 207
654 ACCTGCCATACCATGACGCTGGCGGCAAAATGCGACAGCTCAAGGCG 703
|||:|||||:
207 sProAlaMetThrAlaSerAlaIleAlaLysIleAla.....LeuGln 222
704 TGAACCCCTGTTTCTGCTCGCAACGCTG...CTGCGGACAGAGT 750
||:|||||:
222 yLysTyrProIleIleProCysGlnIleIleArgThrLysGlySerTyr 238
751 TTCGATTTCACATCCGCGCC...GTCCAGGGGATGTGACGGCGACAA 797
||:|||||:
239 PheLysValIleValHisProGlnLeuLysLeuGlnGlnThrGlyAspAs 255
798 AGCCCATGATGCC.....GCCGTGTCAACCGCAATGCCGAT 835
|.....|.....|.....|.....|.....|.....|.....|.....|
255 nLysAlaAspCysTyrAsnIleMetLeuAsnIleAsnIleMetLeuGly 272
836 ATTCGATACGCGCGTTTCCGAGCGAGTATCTGTTATGTACACCGTAC 885
||:|||||:
272 lLysPheLysGlnAsnProAlaGlnThrPheThrPheHisAsnArgTrp 288
886 AAA 888
289 Lys 289

```

seq_name: sp_bacteria:098149

```

seq_documentation_block:
ID 098149 PRELIMINARY; PRT; 328 AA.
AC 098149;
DT 01-OCT-2001 (TREMblrel. 18, Created)
DT 01-OCT-2001 (TREMblrel. 18, Last sequence update)
DT 01-OCT-2001 (TREMblrel. 18, Last annotation update)
DE LIPID A BIOSYNTHESIS LAUROYL ACYLTRANSFERASE.
MR1179.
GN Rhizobium loti (Mesorhizobium loti).
OS Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Phyllobacteriaceae; Mesorhizobium.
OX NCBI_Taxid=381;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MAFP303099;
RX MEDLINE=21082930; PubMed=11214968;
RA Kaneko T., Nakamura Y., Sato S., Asanizu E., Kato T., Sasamoto S.,
RA Watanabe A., Ideasa K., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
RA Mochizuki Y., Nakayama S., Nakazaki N., Shimo S., Sugimoto M.,
RA Takeuchi C., Yamada M., Tadada S.
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
Mesorhizobium loti."
RL DNA Res. 7:331-338(2000).
DR EMBL: AP002996; BAB4814.1;
KW Transferase; Acyltransferase; Complete proteome.
SQ SEQUENCE 328 AA; 37043 MW; 64A39DC394F5E700 CRC64;

```

alignment_scores:
 Quality: 153.00 Length: 324
 Ratio: 0.922 Gaps: 12
 Percent Similarity: 51.235 Percent Identity: 22.840

alignment_block:

US-09-303-518d-569 x 098L49 ..

Align seg 1/1 to: 098L49 from: 1 to: 328

```

10 TTACATTCAGGCGTGTTCCTCCCTTGGGAACCGCATGACATCCCTGT 59
   |||||
   9 LeuArgPheArgTyrGlyArgGlnLeuArgGlnLeuAsnTyrTrpLeuVa 25
   :
   60 GACC.....GCCCTGCAATAGCCCTCCCTGCTGCTGCTGCTT 97
   :
   25 LAlaArgAlaLeuMetIleIleIleSerValLeuArgLeuLeuProValA 42
   :
   98 CCTGTGTCACACGCTGGGAACCGCTCGACATCGCGCTTTTACCTT 147
   :
   42 spSerAlaLeuAsnPheAlaAspArgValAlaArg.....PheIleGly 56
   :
   148 TTAAGGAGACACGCGCGCGCATC...GTGCGCAATATGCTGACGCG 194
   |||
   57 ProLysValGlyArgHisGlnValAlaIleGlyAsnLeuArgLysAlaTy 73
   :
   195 CATGAATCCGACCCCAAAAGCGTCAAGCGCTTTTGGGGAACGCGCA 244
   :
   73 rProGlnLysSerGlnAspGlnIleGlnAlaIleAlaSerAspMetTrpG 90
   :
   245 AAGCGGCTTGGAACTGCCCCCGCGCTTTTC..... 276
   :
   90 LysnMetAlaArgLeuAlaAlaGlnIlePheLeuAspAlaLeuPhe 106
   :
   277 .....AGAAACCGGGAACATGAGAACAAATGTCTCA 308
   |||||
   107 AspTyrAspProAlaAlaSerLysProGlyArgValGln..... 119
   :
   309 AGGCGTACACGCGTGGGAACAT...GTGACAGACGCTTGGCAAAACAG 355
   |||||
   120 ...ValLysGlyIleGlnHisPheValGlnIleAlaGlnLysGlnP 135
   :
   356 AAGCGCTGCTATTCATGACGCGCATGAGTACGATTTGGCGCGGA 405
   :
   135 roHisIleValPhe...ThrGlnHisLeuGlyAsnPheGlnLeu..... 148
   :
   406 CGGTACATACGCCAGCAGCTTCG.....TTCCG 434
   |||||
   149 .....LeuProValAlaAlaIleAlaIleThrPheGlyMetAs 159
   :
   435 GGTGACCGCGCATGTACAAACCGCGGAATCAAGCGTACAAATCA 484
   :
   159 nIleThrAlaLeuPheArgProProAsnAsnProTyrLeuAlaAspTyrI 176
   :
   485 TGCAGCGCGGACAGGTTGCGCGCAAGGAAACCGCGCTACACGATA 534
   :
   176 IeLeuSerThrArgSerThrMetGlyAlaLeuLeuProSerMetAla 192
   :
   535 CAAGGCGTCAAAACAAATCATCAAAACCGCTGCGGCGGCAAGCAACAT 584
   :
   193 GlyAlaSerPheAlaLeuAlaGlyValLeuGlnAsnGlyLysnIleG 209
   :
   585 CGTCTGCCGACGACGCTCCCTCCCTCAAGAGCGCGGGAAGGCGCTAT 634
   :
   209 yAlaLeuValasp.....GlnLysPheSerAsnGlyLeuA 221
   :
   635 GGGTGGATTTCTTGGCAACCTGCTATACATGACGCTGCGGCGGCAAA 684
   :
   221 spTrpThrPhePheGlnArgProGlnSerAsnArgValLeuGlyThr 237
   :
   685 TTGGCACACGTCAAAGCGTGAACACCTGTTTCTGCTGCCAAGCGCT 734
   :
   238 LeuAlaArgHisTyrAspCysAspValTyrProAlaArgLysIleArg 254
   :
   735 GCGTGGCGGACAGGTTGATTTGACATCCG.....C 769
   :
   254 uProGlyAsnArg...PheArgLeuGlnIleGlnAspArgLeuValLeuP 270

```

seq_name: sp_bacteria:082937

```

770 CCGTCAAGGGGAATTGACGCGACAAAGCCCATGATGCCCGCGCTTC 819
   ||
   270 roArgThrAlaAspGlySerValAspAlaHisAlaThrTrpGlnMetLeu 286
   :
   820 AACCGCAATGCCGAATATGGATACGCGCTTTCCGACGACGATGTGTT 869
   :
   287 AsnAspValValGlnArgTrpValArgGlnAspProGlyGlnTrpMetTr 303
   :
   870 TATGACACCGCTACAAATG 891
   :
   303 pheHisLysArgTrpGlnIle 310

```

seq_documentation_block:

```

ID 082937 PRELIMINARY: PRT: 332 AA.
AC 082937:
DT 01-NOV-1998 (TREMBLrel. 08, Created)
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DE 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE LIPID A BIOSYNTHESIS (KDO)2-(LAUROYL)-LIPID IVA ACYLTRANSFERASE.
GN ECF4 OR MSB.
OS Escherichia coli O157:H7.
OC Plasmid pO157.
OC Bacteria: Proteobacteria: gamma subdivision: Enterobacteriaceae;
OX NCBI_Taxid=83334;
RN (1)
RP SEQUENCE FROM N.A.
RX MEDLINE=98290540; PubMed=9596716;
RA Makino K., Ishii K., Yasunaga T., Hattori M., Yokoyama K.,
RA Yatsudo H.C., Kubota Y., Yamachi Y., Iida T., Yamamoto K., Honda T.,
RA Han C., Ohtsubo A., Kasamatsu M., Hayashi T., Kuwara S., Shinagawa H.,
RT "Complete nucleotide sequences of 93-kb and 3.3-kb plasmids of an
RT enterohemorrhagic Escherichia coli O157:H7 derived from Sakai
RT outbreak."
RT DNA Res. 5:1-9(1998).
RN (2)
RP SEQUENCE OF 1-235 FROM N.A.
RX STRAIN=4304-PT14;
RA Boettlin P., Chen S., Colbourne J.K., Johnson R., De Grandis S.,
RA Gyles C.,
RT "Evolution of enterohemorrhagic Escherichia coli hemolysin plasmids
RT and the locus for enterocyte effacement in shiga toxin-producing E.
RT coli."
RL Infect. Immun. 66:2553-2561(1998).
RN (3)
RP SEQUENCE FROM N.A.
RX STRAIN=EDL933;
RA Brunder W.,
RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB011549; BAA31840.1; -.
DR EMBL: AF043470; AAC24348.1; -.
DR EMBL: Y11275; CAA72141.1; -.
KW transferase; Acyltransferase; Plasmid.
SQ SOURCE 332 AA; 37858 MW; 0D49F6B93E29E9A6 CRC64;

```

alignment_scores:

```

Quality: 151.00 Length: 286
Ratio: 0.993 Gaps: 9
Percent Similarity: 53.147 Percent Identity: 22.028

```

alignment_block:

US-09-303-518d-569 x 082937 ..

Align seg 1/1 to: 082937 from: 1 to: 332

```

82 TCCTGTGCGCGCTTCCCTGTCGACACGCTGGGAACCGCGCTGCGACA 131
   :
   56 AlaMetValProProAlaLeuArgAspProLeuLeuGlyLysLeuGlyme 72

```



```

549 AATCATCAAGCCCTGCTTCGGCGAAGCAACATCGTCCTGCC... 594
223 rPheValArgSerValArgGlnGlyTyrTrpGlyTyrTyrLeuProAspG 240
595 ..GACCAGCTCCCTCCCTCCCTCAAGAGCGGGGAAGCGGTATGGTGGAT 642
240 LAspHisGlyProGlnPheSerGlu.....PheAlaAsp 251
643 TTCTTCGGCAACCTGCTATACATGACGCGCGCAAAATGGCACA 692
252 PhePheAlaThrTyrLysAlaThrLeuProValIleGlyArgLeuSerAr 268
693 CGTCAAGAGCGGTGAACACCTGTTTCTGCGCAACGCTCGCTGCGC 742
268 gIleSerGlyAlaArgIleIle.....ProLeuPheProValT 281
743 GACAAGGT.....TTCGATTGGCACATCCGCCCGTCACAGG 780
281 YrAspGlyLysThrHisHisLeuThrIleHisValSerProProLeuAla 297
781 GAATTGAACGGCGAAGCAACCATGATGCGCGCTGTCAACGGCAATGC 830
298 IleArgGlnLysSerAspAlaHisIleAlaArgGlnIleAsnGluValVa 314
831 CGAATTATGATACGCGCTTTCCGACGCAATCTGTTATGTATACACC 880
314 IGlAsnPheValArgProHisProGlnGlnTyrThrTrpIleLeuLysL 331
881 GCTACAAA 888
331 euleuLys 333

```

seq_name: sp_bacteriap:Q92GN2

```

seq_documentation_block:
ID Q92GN2 PRELIMINARY; PRT; 290 AA.
AC Q92GN2;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DE 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE LIPID A BIOSYNTHESIS LAUROYL ACYLTRANSFERASE.
GN HTRB OR RC1091.
OC Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
OC Rickettsiaceae; Rickettsia; Rickettsia.
OX NCBI_TaxID=781;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MALISH 7;
RX MEDLINE=21442074; PubMed=11557893;
RA Ogata H., Audic S., Renesto-Audiffren P., Fournier P.-E., Barbe V.,
RA Samson D., Roux V., Cossart P., Weissenbach J., Claverie J.-M.,
RA Raoult D.;
RT "Mechanisms of evolution in Rickettsia conorii and R. prowazekii.";
RL Science 293:2093-2098(2001).
DR EMBL; AE008660; ALU03629.1;
KW Transferase; Acyltransferase; Complete proteome.
SO SEQUENCE 290 AA; 33522 MW; 1C0PD05318332ED0 CRC64;

```

alignment_scores:

Quality:	149.00	Length:	201
Ratio:	1.221	Gaps:	8
Percent Similarity:	60.697	Percent Identity:	23.383

alignment_block:

US-09-303-518D-569 x Q92GN2

Align seg 1/1 to: Q92GN2 from: 1 to: 290

```

310 GCGGTACGCGCTGGAGACATGTGCACAGCGCTTGGACAAACAGAGG 359
||||| ||| |||:||||| ||| |||

```

```

101 AlaIleIleGlyIleGluAsnIleLys.....LysLeuGluG 113
360 G.....CGCTATTCATACACCGCCGACATCGGACCTAGCATTTGGCG 403
113 yGlnProPheLeuLeuPheGlyGlyHisPheAlaAsnTrpAspIleSer 130
404 GACGCTACATCAGCCAGCAGCTCCGTTCCCGTCAGCGCGCATGTACAA 453
130 IleLysIleLeuHisLysPheTyrPro...LysValAlaValIleTyrArg 145
454 CGCCGGAATATCAAGCATAGACAAATCTGACAGCGGGGAGGCTTCG 503
146 LysAlaAsnAsnProTyrValAsnLysLeuValAsnLysSerArgAlaG 162
504 CGGCAAGAAAAACCGCGCTTACCATCATACAAAGGGGTCAACCAATCA 553
162 YrAspLysLeuArgLeuIleProLysGlyProGlnGlySerArgAlaLeu 179
554 TCANAAGCCCTGCTCGGGCGAAGCAACATCGTCTGCGCCGACACGTC 603
179 AlaArgAlaIleLysGlnSerGlnSerIleValMetLeuValAsp..... 193
604 CCTCCCTCCATCAAGAGCGGGGAAGCGGTATGGGTGATTTCTTGGCAA 653
194 .....GlnLysMetAsnAspGlyIleGluValProPheLeuGlyH 207
654 ACCTGCTATACATGACGCTGCGCAAAATGGACACGTCACAAAGGCG 703
207 sProAlaMetThrAlaSerAlaIleAlaLysIleAla.....LeuGlnT 222
704 TGAAGACCTGTTTCTGCTGCGAAGCGCTG...CTTGGCGACAAAGT 750
222 YrLysTyrProIleIleProCysGlnIleIleArgThrLysGlySerYr 238
751 TTCGATTGGACATCCGCC...GTCAAGGGGAATTGAAGGGGACAA 797
239 PheLysValIleValHisProGlnLeuLysPheGlnGlnThrGlyAspAs 255
798 AGCCCATGATGCC.....GCCGTGTCAACCGCAATGCCGAT 835
255 nLysAlaAspCysTyrAsnIleMetLeuAsnIleAsnGlnMetLeuGly 272
836 ATTGATACGCGCTTCCGACGATATCTGTTATATACACCGCTAC 885
272 LutrPallValLysGlnAsnProAlaGlnTrpPheTrpPheHisAsnArgTrp 288
886 AAA 888
289 Lys 289

```

seq_name: sp_bacteria:P94909

```

seq_documentation_block:
ID P94909 PRELIMINARY; PRT; 529 AA.
AC P94909;
DT 01-MAY-1997 (TREMBLrel. 03, Created)
DT 01-MAY-1997 (TREMBLrel. 03, Last sequence update)
DT 01-NOV-1998 (TREMBLrel. 08, Last annotation update)
DE HYPOTHETICAL 58.7 KDA PROTEIN (FRAGMENT).
OS Microbacterium ammoniaphilum.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Micrococccineae; Microbacteriaceae; Microbacterium.
OX NCBI_TaxID=42460;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 15354;
RX MEDLINE=96257250; PubMed=8654988;
RA Striebel H.M., Seebor S., Jarsch M., Kessler C.;
RT "Cloning and characterization of the Mann restriction-modification
RT system from Microbacterium ammoniaphilum in Escherichia coli.";
RL Gene 172:41-46(1996).
DR EMBL; X79027; CAA55649.1;
KW Hypothetical protein.

```

FT NON_TER 1 1
 50 SEQUENCE 529 AA: 58716 MW: 5DC1216237DEF772D CRC64:

alignment_scores:
 Quality: 147.50 Length: 293
 Ratio: 1.209 Gaps: 16
 Percent Similarity: 41.638 Percent Identity: 28.669

alignment_block:
 US-09-303-518D-569 x P94909 ..

Align seg 1/1 to: P94909 from: 1 to: 529

```

30 CCCTTGGCAACCGCATGACATCTGTTGACCGCGCTGCAATGCC 79
   ||| ..... ||| ||| ||| ||| ||| ||| ||| ||| |||
107 rArGgIuPrOAlAaGrPrOAlAlAaGrAsPaLaArGrAlhIsLeuArG 123
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
74 PrOAlGInArGrAlhIsPrOleuGInArGrPrOHISGLuArG 90
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
80 T..... 80
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
90 oGlyAlAaGrGlyGluLeuPrOAlAaGrAsPaLaArGrAlhIsLeuArG 107
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
81 .....CTCCGCTGCGCGCTTCTGCTGTCACAC..... 110
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
107 rArGgIuPrOAlAaGrPrOAlAlAaGrAsPaLaArGrAlhIsLeuArG 123
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
111 ...GCTGGAAACCGCGTCGACATCTGCGCTTTAACCTTTAAAGGAAG 157
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
124 LeuGlyAlhIsPrOAlAaGr...ArGlyValLeu..... 134
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
158 ACCCGCGCGCATGCGCGCAATGCGTCAGCGGACGATGATCCGAC 207
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
135 ...ArGrAlhIsArGrAlhIsArGrAlAaGrGlyGluArGlyArG 150
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
208 CCCAAAGCGGTCAAGCGCTTTTGGCGAAACGCGCAAAAG...CGCTTT 254
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
150 rOArGgInGInAlPrOArGInhIsPrOArGlyArGrAsPaLaArG 166
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
255 GGAAGTTCGCCCGCGCTTTTCAGAAACCGGAAAGCATGAAACATGT 304
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
167 Gly.....ArGrPrOAlhIsArGrAlhIsArGrAlhIsArGrAlh 176
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
305 TCAGAGCGGTACAGCGGTGGAAACATGTGACAGAGCTTTCAGAAAC 354
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
176 gArGrAlAaGrArGrAlhIsArGrAlhIsArGrAlhIsArGrAlh 193
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
355 GAAGGCGCTGATTTCATGACGCGCGACATGCGGACGATTTGGCGG 404
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
193 In.....hIsArGrAlAaPrOArGrAlhIsArGrAlhIsArGrAlh 205
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
405 AGCTACATCAGCGAGCGCTTCGCTCCGCGTACCGGACATGTACAAAC 454
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
206 AlAaGlyhIsPrOleu..... 211
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
455 CGCGCAAAATCAAGCGATAGCAAAATCATGACGCGGACGCGTCCG 504
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
212 .....ArgGlyGlyAlhIsPrOleu..... 219
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
505 GCGCAAGGAAACCGCGCTAC.....CAGCATACAGGCGGTCAACA 548
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
219 lAAlAaGrPrOleuArGrAlhIsArGrAlhIsArGrAlhIsArGrAlh 235
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
549 AATCATCAAGCGCTGCTCGG..... 572
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
236 AlAaGlyArGrAlhIsPrOAlhIsArGrAlhIsArGrAlhIsArGrAlh 252
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
573 .....CGAAGCAACCATGCTGCTCGCGGACGCGGACGCTCCCGCA 615
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
252 nGlyAlAaGrAlAaAlhIsArGrAlhIsArGrAlhIsArGrAlhIsArGrAlh 268
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
616 GAAGGCGGGAAGG.....CGTATGCGGATTTCTTCGCGCA 653
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
268 rArGrAlhIsArGrAlhIsArGrAlhIsArGrAlhIsArGrAlhIsArGrAlh 284
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

654 ACCTGCGCTATACATGAAGCGTGGCGCAAAATTCGACAGCTCAAGCG 703
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
285 AlA.....ArGrAlAaGlyGly.....GlyGlyLeuPrOArGrAlh 296
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
704 TGAAGACCGCTGTTTCTGTCGCGAAGCGCTGCGCGGACAGCTTTC 753
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
296 gGluAsp.....ArgValArG..... 301
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
754 GATTTCACATCGCGCGCGTCCAGGGA 782
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
302 .....AlhIsArGrPrOArGrPrOArGly 309
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

seq_name: sp_bacteria:Q9AKK5

seq_documentation_block:
 ID Q9AKK5 PRELIMINARY; PRT; 290 AA.

```

AC Q9AKK5;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE HTB PROTEIN.
GN HTB.
OS Rickettsia montana.
OC Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
OC Rickettsiaceae; Rickettsiae; Rickettsia.
OX NCBI_TaxID=33991;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21219194; PubMed=11319266;
RA Andersson J.O., Andersson S.G.E.;
RT "Pseudogenes, junk DNA and the dynamics of Rickettsia genomes.";
RL MOL. Biol. Evol. 18:829-839(2001).
DR EMBL; AJ293330; CAC33650.1; ..
SQ SEQUENCE 290 AA; 33635 MW; 9599B3C0C3C076F9 CRC64;

```

alignment_scores:
 Quality: 147.00 Length: 190
 Ratio: 1.267 Gaps: 7
 Percent Similarity: 61.053 Percent Identity: 23.158

alignment_block:
 US-09-303-518D-569 x Q9AKK5 ..

Align seg 1/1 to: Q9AKK5 from: 1 to: 290

```

343 TTGGACAACACGAAGG.....CTGCTATTTCATCAGCGCGACATCG 386
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
108 lLeuYsLeuGInGlyGInPrOphelLeuPheserGlyhIsPhel 124
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
387 CAGCTACGATTTGGCGGCGCTACATCAGCGAGCGCTTCGCTCCGCG 436
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
124 aAsnTrpAspIleSerLeuYsLeuYsLeuYsPhelTyPro..LysV 140
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
437 TGACCGCATGTACAAACCGCGCAAAATCAAGCGATAGCAAAATCATG 486
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
140 AlAlAaValIleTyArGlyAlAaIleYsLeuYsLeuYsLeuVal 156
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
487 CAGCGGCGAGGCTTCGCGCAAGGAAACCGCGCTCCGCGCATAC 536
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
157 AsnGlySerArGAlhIsArGrAlhIsArGrAlhIsArGrAlhIsArGrAlh 173
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
537 AGGGGTCAAGCAAAATCATCAAGCGCTGCTGCGGCGAAGCAACATCG 586
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
173 uGlySerArGAlAaValArGAlAaIleYsLeuYsLeuYsLeuValM 190
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
587 TCCTCCCGGACGACGCTCCCTCCCTCAAGAGCGGGAAGGCGTATG 636
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
190 eLeuValAsp.....GlnYsMetAsnAspGlyIleGlu 201
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
637 GTGATTTCTTGGCAACCTGCGTATTCATGACGCTGCGGCAAAAT 686
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```


seq_documentation block:
 ID 092983 PRELIMINARY: PRT: 467 AA.
 AC 092983;
 DT 01-MAY-1999 (TREMBLrel. 10, Created)
 DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE ACYLTRANSFERASE.
 GN HTRB OR CPN0098.
 OS Chlamydia pneumoniae (Chlamydia pneumoniae).
 OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydiales.
 OX NCBI_TaxID=83558;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CWL029;
 RC MEDLINE=99206606; PubMed=10192388;
 RA Kalman S., Mitchell W., Marathe R., Lammel C., Fan J., Hyman R.W.,
 RA Olinger L., Grimwood J., Davis R.W., Stephens R.S.;
 RT "Comparative genomes of Chlamydia pneumoniae and C. trachomatis";
 RL Nat. Genet. 21:385-389(1999).
 DR EMBL: AF001596; AAD18251.1; -;
 KW Transferase; Complete proteome.
 SQ SEQUENCE 467 AA; 53193 MW; D3C7C284E9220DD0 CRC64;

alignment_scores:
 Quality: 146.00 Length: 331
 Ratio: 0.869 Gaps: 13
 Percent Similarity: 50.755 Percent Identity: 22.356

alignment_block:
 US-09-303-518D-569 x 092983 ..

Align seg 1/1 to: 092983 from: 1 to: 467

```

34 TTGGGAACCGCCGATGACATCTCTTGACCGCCCTGCTCAATGCTCTC 83
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
13 LenglualAProleuTYrLeuValSerGIylelleAlaLeuCYsAr 29
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
84 CCTGTGCGGCTTCTGCTGACACGCGTGGAAACCGGCTCGACATC 133
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
29 GHISrThProArSerPheLeuThrGlyLeuGlyLysGlyPheGlyPhe 46
134 TGGCGTTTACCTTTTAAGAAGACGCGCGGCATCGTCGCAATATG 183
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
46 euAlaPheTYrIleleSerAspTYrArgLYsThrAlaLeuThrAsnLeu 62
184 .....CG 185
63 AlaLeuAlaPheProGluLysThrPheAspGluArgTYrLysIleAlaAr 79
186 TCAGCGAGGATGATCCCGACCCCAAAACGGTCAAGCGCTTTTGGCG 235
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
79 glnIserLeuGlnIleAsnIleAspLYsLeuIleThrLeuLeuAlaIleG 96
236 AAACGGCAAAAGCGGTTTGAACCTGCCCGCGCTTTTC..... 276
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
96 lngIleuValGlyAsnIleAspLYsLeuIleThrIleValThrSerSer 112
277 AGAAAAACCG.....GAAGACAT 293
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
113 ArgAsnProLYsGlyPheSerSerGluValIleSerAsnGluAspLe 129
294 AGAAACATGTTCAAGCGGTACACGCGTGGAGACATGTGACGAGGCTT 343
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
129 uGluGluThrPheLYs.....AsnL 136
344 TGGACAAACGAGAGGCTGATATCATACGCGGACATCGGACATAC 393
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
136 euGlnIleLYsGlnGlyLeuIleLeuPheCYsGlyHisIleAlaSnrTP 152
394 GATTGGGGGAGCGTACATCAGCCAGCAGCTTCGTTCCCGCTGACCGC 443
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

153 GluLeuProPheLeuTYrIleThrLYsAsn.....TYrProGlyIleAl 167
444 CATGTACAAACCGCCGAATCAAGCATACGAAATCATGACGCGG 493
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
167 aPheAlaIleValIleLYsAsnGlnArgLeuSerLYsIleIleAla 184
494 GCAGGGTTCCGGGAAAGAAACCGCGCTACCATCATACAGGGGTC 543
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
184 euArgGluValPheLYsGlyLYsIleValProLYs.....AsnGlyIle 199
544 AAACAAATCATCAAAAGCGGTTGGGCGAGACAAACCATGCTGCC 593
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
200 GlnGlnGlyIleGlnAlaLeuAsnGlnGlyLYsLeuValGlyIleVal 216
594 CGACCAGCTC.....CCGCCCCCTCAAGAG 619
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
216 yAspIleAlaLeuLeuMetSerSerTYrThrPro..... 228
620 GCGGGGAGCGGTATGGTGATTTCTTGCGCAACCTGCTATACATG 669
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
229 .....LeuPheGlySerProAlaPheThr 237
670 ACGTGGCGGCAAAATTTGACACGTCAAGCGGTGAAACCGCTTTT 719
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
238 ThrSerProAlaLeuLeuAlaTYrLYsThrGlyPheProValIleAla 254
720 CTGTGCGCAACCGCTGCGTGGGCGACAGTTTGATTTGCACATCCG 769
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
254 lAsnValSerArg.....GlnAlaLYsGlyPheGlu.....ValIleP 267
770 CCGTCCAAAGGAGATTTGAACGCGCAAAAGCC.....CATGAT 807
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
267 ro...SerAlaLYsLeuTYrAlaAsnLYsSerLeuProMetLYsGluSer 282
808 GCGCGCGTGTCAACCGCAT.....GCCAATATTGATGATCG 845
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
283 ValAlaIleLeuMetAspGlnMetMetGlyPheLeuGlyLYsGlyIleAl 299
846 CCGTTTTCGACGATATCTGTTATGTACACCGCTACAA 888
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
299 aSerGlnProGluGlnIleThrMetIleIleLYsAsnArgTYrLYs 313

```

seq_name: sp_bacteriap:09KVD4

```

seq_documentation block:
ID 09KVD4 PRELIMINARY: PRT: 325 AA.
AC 09KVD4;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE LIPID A BIOSYNTHESIS (KDO)2-(LAUROYL)-LIPID IVA ACYLTRANSFERASE.
GN VC0212.
OS Vibrio cholerae.
OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
OX NCBI_TaxID=666;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=EL TOR N16961 / SEROTYPE O1;
RC MEDLINE=20406833; PubMed=10952301;
RA Heidelberg J.F., Harte D.H., Hickey E.K., Peterson J.D., Umayam L.A.,
RA Dodson R.J., Haft D.H., Read T.D., Retlein H., Richardson D.,
RA Gill S.R., Nelson K.E., Reed T.D., Tettein H., Dragoti I., Sellers P.,
RA Ermolaeva M.D., Yamachyan J., Bass S., Qin H., Dragoti I., Sellers P.,
RA McDonald L., Utterback T., Fleischmann R.D., Niernan W.C., White O.,
RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
RA Fraser C.M.;
RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio
RT cholerae";
RT Nature 406:477-483(2000).
RL EMBL: AE004111; AAF93388.1; -;
DR TIGR: VC0212; -;
KW Complete proteome.
SQ SEQUENCE 325 AA; 37215 MW; 3F949ED94B5C5F92 CRC64;

```

alignment_scores:

Quality: 141.00 Length: 186
Ratio: 1.270 Gaps: 6
Percent Similarity: 59.677 Percent Identity: 25.269

alignment_block:

US-09-303-518D-569 x Q9KVD4 ..

Align seg 1/1 to: Q9KVD4 from: 1 to: 325

```

319 GGCCTGGGACATGTCAGCAGCGCTTTGGACAAACAGCAGGCTGCTATT 368
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
122 GtlyglulnsluLeuProLeuLeuaspserGlygluArgValIlele 138
    : : : : : : : : : : : : : : : : : : : : : : : : :
369 CATCACCGCCACATCGCAGCTACGATTGGGCGACGCTACATCAGCC 418
    : : : : : : : : : : : : : : : : : : : : : : : : :
138 uluValProHisserTrpAlaIleAspTyrAlaAlaValMetIleuAla 155
    : : : : : : : : : : : : : : : : : : : : : : : : :
419 AGCAGCTCCGCTCCGCTGACCGCCATGACAAACCGCGAAATCAAA 468
    : : : : : : : : : : : : : : : : : : : : : : : : :
155 laArg...GlyTyrLysValAlaAlaAsnIleMetLysProGlnArgAsnPro 170
    : : : : : : : : : : : : : : : : : : : : : : : : :
469 GCGATACACAAATCATGACGCGGCGAGGCTCGCGGCAAGCAAAAC 518
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
171 lIleAlaAspTrpLeuMetHisValAlaArgMetGlnTyrGlyArgIle 187
    : : : : : : : : : : : : : : : : : : : : : : : : :
519 CGCGCTTACCGACATACAGAGGCTCAACAAATCATCAAGCCCTGCT 568
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
187 ePhe...ThrArgIleSerGlyIleLysProPheLeuArgSerIleGln 203
    : : : : : : : : : : : : : : : : : : : : : : : : :
569 CGGCGCAAGCAACCATCGTCCTCCGCGACACGCTCCCTCCCTCAAG 618
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
203 erGlyTyrValGlyTyrTrpValProAspGluAspHisGlyProGln 218
    : : : : : : : : : : : : : : : : : : : : : : : : :
619 GCGGCGAAGCGCTATGGGTGATTTCTCGGCAACCTGCTATTCAT 668
    : : : : : : : : : : : : : : : : : : : : : : : : :
219 .....AsnSerValPheValProPhePheAlaThrGluAlaThrle 233
    : : : : : : : : : : : : : : : : : : : : : : : : :
669 GACGCTGGCGGCAAAATGGCACACGTC...AAAGCGTGAAACCCCTG 715
    : : : : : : : : : : : : : : : : : : : : : : : : :
233 ulysGlyPheGlyLysMetAlaLysLysCysAlaHisValProL 250
    : : : : : : : : : : : : : : : : : : : : : : : : :
716 TTTTCTGCTGCGAACGCTGCTCGGCGGCAAGTTTGCATTTGACATC 765
    : : : : : : : : : : : : : : : : : : : : : : : : :
250 eumetSerCysTyrAsnSerAspSerGlyArg...TyrGluValHisIle 265
    : : : : : : : : : : : : : : : : : : : : : : : : :
766 CCGCCC...GTCCAAGGGAATTGAAGCGGCAAAAGCCCATGATCGGC 812
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
266 LeuProAlaLeuGlnAsnProThrGlyAspGluGluAlaAspAlaLe 282
    : : : : : : : : : : : : : : : : : : : : : : : : :
813 CGGTTCAGACGCAATGCCGATATTGATACGCGCTTTCCGACGCGAT 862
    : : : : : : : : : : : : : : : : : : : : : : : : :
282 uAlaMetAsnArgAlaIleGluAlaLeuValThrProGlnProGluGln 299
    : : : : : : : : : : : : : : : : : : : : : : : : :
863 ATCTGTTT 870
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
299 yrmEtrp 301
    : : : : : : : : : : : : : : : : : : : : : : : : :
seq_name: sp_bacteriap:Q9KVD3

```

seq_documentation_block:
ID Q9KVD3 PRELIMINARY; PRT; 318 AA.

AC Q9KVD3;
DT 01-OCT-2000 (TREMblrel. 15, Created)
DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)
DE LIPID A BIOSYNTHESIS LAUROYL ACYLTRANSFERASE.
GN VC0213.
OS Vibrrio cholerae.
OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrrio.
NCBI_TaxID=666;

```

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=EL TOR N16961 / SEROTYPE O1;
RX MEDLINE=20406833; PubMed=10952301;
RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwinn M.L.,
RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,
RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
RA Ermolaeva M.D., Vamathevan J., Bass S., Olin H., Dragoi I., Sellers P.,
RA McDonald L., Uitterlind T., Fleischmann R.D., Nierman W.C., White O.,
RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
RA Fraser C.M.;
RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio
RT cholerae."
RL Nature 406:477-483(2000).
DR EMBL; AE004111; AAF93389.1; -
DR TIGR; VC0213; -
RW Transferase; Acyltransferase; Complete proteome.
SQ SEQUENCE 318 AA; 36542 MW; FE95D7AC83106E1 CRC64;

```

alignment_scores:

Quality: 136.50 Length: 317
Ratio: 0.858 Gaps: 15
Percent Similarity: 50.158 Percent Identity: 24.921

alignment_block:

US-09-303-518D-569 x Q9KVD3 ..

Align seg 1/1 to: Q9KVD3 from: 1 to: 318

```

58 TTGACCGCCCTGCTCAAAATGCTTCCTGCTGCGGCTTCTGCTGCA 107
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
32 LeuLeuAlaLeuLeu.....ValAsnLeuLeuProTyrProValIlele 46
    : : : : : : : : : : : : : : : : : : : : : : : : :
108 CACGCTGGGAACCGGCTCGCA..... 129
    : : : : : : : : : : : : : : : : : : : : : : : : :
46 ulysIleGlyArgGlyLeuGlyGlnPheSerMetArgPheGlyLysIys 63
    : : : : : : : : : : : : : : : : : : : : : : : : :
130 .....CATCTGGGCTTTTACTTTTAAAG 153
    : : : : : : : : : : : : : : : : : : : : : : : : :
63 rGValHisIleAlaArgArgAsnLeuGluLeuAlaPheProThrMetSer 79
    : : : : : : : : : : : : : : : : : : : : : : : : :
154 GACAGACGCG.....GCGCGCATGCTGCGCAATATGCGTCAGGAGCAT 197
    : : : : : : : : : : : : : : : : : : : : : : : : :
80 GlnSerGluIleAspAlaPheValIleuGluAsnPhelysnThrGly.. 95
    : : : : : : : : : : : : : : : : : : : : : : : : :
198 GAATCCGCAACCCCAAAACGGTCAAAGCGCTTTTGGCGAAAGCGCAAA 247
    : : : : : : : : : : : : : : : : : : : : : : : : :
96 .....AlaAlaLeuIleGluThrGlyIleT 104
    : : : : : : : : : : : : : : : : : : : : : : : : :
248 GCGGTTTGAACCTTCCCCCGGCTTTTTCAGAAACCGGAACATATGAA 297
    : : : : : : : : : : : : : : : : : : : : : : : : :
104 hTrpPheTrpProThrTrpArgPheLysArgIleLeuIleAspLysAsp 120
    : : : : : : : : : : : : : : : : : : : : : : : : :
298 ACAATGTTCAAAGCGGTACACGCGGTGGGAACATGTGCAGCAGGCTTGA 347
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
121 Thr.....GlnAlaIleArg 125
    : : : : : : : : : : : : : : : : : : : : : : : : :
348 CAACAC.....GAAGGCGCTATTCATCAGCCCGCCACATCG 385
    : : : : : : : : : : : : : : : : : : : : : : : : :
125 gGlnHisAlaLysThrGlyGlnGlyValIleuLeuCysCysValHisAla 142
    : : : : : : : : : : : : : : : : : : : : : : : : :
386 GCACTACGATTTGGGCGGACGCTACATCGCCAGAGCTTCGCTCCCG 435
    : : : : : : : : : : : : : : : : : : : : : : : : :
142 euAsnLeuGluIleThrAlaArgAlaPheAla...ValIleuGlyIleGly 157
    : : : : : : : : : : : : : : : : : : : : : : : : :
436 CTGACGCGCATGTTCAAACCGCGCAAAATCAAGCGCATATGACAAATCAT 485
    : : : : : : : : : : : : : : : : : : : : : : : : :
158 GlyTyrGlyValTyrArgProHisSerAsnProAlaTyrGluPheIleG 174
    : : : : : : : : : : : : : : : : : : : : : : : : :
486 GCAGGCGGCGAGGCTTCGCGCAAGCAAAACCGGCTACACAGATAC 535
    : : : : : : : : : : : : : : : : : : : : : : : : :
174 nTyrArgGlyArgThrArgAsnGlyAsnGlnLeuIleAsnArgThr.... 189
    : : : : : : : : : : : : : : : : : : : : : : : : :

```

100

EMBL: D97042; AAB58160.1; -
 DR InterPro: IPR002215; HLYD;
 DR Pfam: PF00529; HLYD; 1
 SO SEQUENCE 405 AA; 44441 MW; CCDC22ECFA35167F CRC64;

alignment_scores:
 Quality: 133.50 Length: 212
 Ratio: 1.362 Gaps: 10
 Percent Similarity: 46.226 Percent Identity: 28.774

alignment_block:
 US-09-303-518D-569 x 006470 ..

Align seg 1/1 to: 006470 from: 1 to: 405

```

150 AAGGAGAGACCGCGCGCATGCTGCCAATATGCTGAGGAGCATGA 199
    ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:
161 LysArgGlnProGlnGlyArgArgSerArgAlaGlyAsnGlyAlaHisGly 177
    ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:
200 ATCCCGACCCCAAAAGCGTCAAAAGCGTTTTCGGAACGCGCAAAAGGC 249
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
177 nserArgLeuTyr.....AlaHisHisA 185
    ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:
250 GGTGTGACACTTGGCCCCGCGTTTTCAGAAACCGGAGACATAGAAC 299
    ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:
185 rgaIagIyValGlyProArgValAlaArg.....GlyAsnHisAlaArg 199
    ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:
300 AATGTCAAAAGCGGTACACGCTGGGAACATGTGA.....GCAGCGTT 343
    ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:
200 GlnArgArgValGlyArgArgValGlyArgAlaAlaAspArgLeuLys 216
    ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:
344 TGA.....CAACACGAAAGGCT.....G 363
    ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:
216 eglValAlaAspLeuArgValAlaArgArgArgArgAlaAspLeuProA 233
    ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:
364 CTATTCATTCAC..... 374
    ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:
233 lalIeHisGlnArgArgAlaGlnArgProGlnGlyAlaArgAlaArg 249
    ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:
375 .....GCCGACATCGGACGCTACGATTTGGCG 403
    ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:
250 ProArgGlnArgAsnArgLeuAlaGlnArgArgArgArgLeuArg 266
    ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:
404 GACGCTACATGACGACGACGCTTCGCGTCCGCGACGCGCATACAAA 453
    ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:
266 g.....GlnProAlaArgHisValVal.....ArgHisAspProA 278
    ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:
454 CCGCGAATAATCAAAAGCGTAGACAAATCATGACGCG..... 491
    ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:
278 rgaIaArgProLeuArgGlnArgGlnArgHisProGlnProGlnProLeu 294
    ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:
492 .....GGCGACGCTTCGCGCAAAAGGAAACCGCGGCTTACGACATAC 535
    ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:
295 ArgThrArgGlnGlyArgGlnArgGlnArgAlaAlaArgGlyAlaAlaArg 311
    ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:
536 AAGGGGTCAAAACATCATCAAAAGCGCTTCGCGCAAGCAACATCATC 585
    ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:
311 gArArgArgGlnArgGlnArgHisArgProGlnGlnGlnValAlaArgVal 328
    ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:
586 GTCTGTCCCGACCGACGTCCTCC..... 608
    ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:
328 rgrProAlaGlyProArgValAlaSerArgSerAlaAlaArgAspAlaAla 344
    ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:
609 .....CCCTCAAGAGGCGCGGGAAGCGGTATGGGT 638
    ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:
345 ArgGlnProAlaArgAspArgGlnArgAlaValGly 356
    ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:

```

seq_name: sp_bacteria:051415

seq_documentation_block:
 ID 051415 PRELIMINARY: PRT: 684 AA.
 AC 051415;

DT 01-NOV-1996 (TREMblrel. 01, Created)
 DT 01-NOV-1996 (TREMblrel. 01, last sequence update)
 DT 01-DEC-2001 (TREMblrel. 19, last annotation update)
 DE ALGINATE LYASE.
 GN ALG.
 OS Pseudomonas aeruginosa.
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
 OC Pseudomonas
 OX NCBI_TaxID=287;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN-A22 ALG+;
 RA Schmitt Andrieu L., Hulen C.;
 RT Cloning and sequencing of a new alginate lyase, algY, from mucoid
 strains of Pseudomonas aeruginosa, and its expression in E. coli.;
 RL Submitted (SEP-1995) to the EMBL/Genbank/DBJ databases.
 DR EMBL; Z54213; CAA9040.1; -
 KW Lyase.
 SO SEQUENCE 684 AA; 74958 MW; CVCCE638F5675D7A CRC64;

alignment_scores:
 Quality: 129.50 Length: 254
 Ratio: 1.028 Gaps: 10
 Percent Similarity: 49.606 Percent Identity: 26.378

alignment_block:
 US-09-303-518D-569 x 051415 ..

Align seg 1/1 to: 051415 from: 1 to: 684

```

42 CGCATGACATCTGTTGAC...CGCCCTGCTCAATAGCTTCGCCGCG 88
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
304 ArgArgGlnHisProHisSerAlaThrProSerAlaThrArgSerProTh 320
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
89 PGCGCGTTTCGTTGCGACACGCTGGGAACCGCGTGGCATCTGGCG 138
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
320 rserSer.....SerGlyLeuAlaAlaArgProAlaGlyArgGlyG 335
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
139 TTTTACCTTTTAAGGAGACCGCGCGGATGCTGCCAATATGCGCA 188
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
335 LysProArgGlnAlaHisProArgGlyGlnArgGlnGlyAlaLeuSer 351
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
189 GCGGACATGANTCCGCGACCCCAAAAGCGTCAAGCGCTTTTGGCGAAA 238
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
352 AlaAlaAlaThrAlaValSerAlaAlaValLeuAlaArgGlnLeuGly.. 367
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
239 CGGCAAAAGCGGTTGGAACCTTGCCTCGCGCTTTTTCAGAAAACCGGAA 288
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
368 .....GlyA 369
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
289 GACATGAAACAATGTTCAAAAGCGGTACACGCGTG.....GGA 326
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
369 rgrProArgThrGlyArgArgThrGlyArgGlnArgProAspLeuGlyArg 385
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
327 ACATGTGCACAGCGTTTGGCAAAACGAGGCGCTATTTCATCAGCG 376
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
386 AlaValGlyGlyGlyGlyGlyGlnAsnArgArgAla.....ProAl 400
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
377 CGCATCGGACGCTACGATTTGGCGGACGCTACATCAGCGGCGAGCT 426
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
400 aSerGlySerProThrAlaTyrGlyGlyGlyAlaArgHisProProIaGlyA 417
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
427 CCGTTCGCGTGAACCGCATGTCAAAACCGCC...GAAATCAAAAGCAT 473
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
417 sPArgProArgAsnArgArgArgSerLeuAlaArgArgGlyGlnAlaAsp 433
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
474 AGACAAATCATGCGCGCGGAGCGGCTTCGCGCAAAAGGAAACCGCGC 523
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
434 ArgThrHisLeuAlaArgArgAsnHisArgArgAlaAlaValLeuArgPrt 450
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
524 CTACCGCATCAAGGCGTCAAAACATCATCAAGCGCTTCGCGGCG 573
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```



```

351 SerProSerProGluArgSerSerThnGlyProGluProProAlaProTh 367
77 GCGCTCCCTGCTGCGCGCTTCTGCTGTCACACAGCGGGAACCG... 122
367 rProLeuLeuAla.....GluArgHisGlyGlySerProGluP 380
123 ....GCTCGGACA.....TCTGGC 137
380 rLeuAlaThrThrProLeuSerGlnGluProValAsnProProSerGlu 396
138 GTTTTACCTTTAAAGAGA.....CCGCGCGCCATCGTCCCAATA 181
397 AlSerProThrArgAspArgSerProProLysSerProGluLysLeuPr 413
182 TGGGTAGCGAGCATGAATCCCGACCCCAAAACGGTCAAGCGCTTTT 231
413 ogInSerSerSerGluSerSerProProSerProGluProThLysV 430
232 GCGGAACGCAAAAGCGGTTTGGAACTTCCCGCGCTTTTTCAGAA 281
430 alSerArgHisAlaSerSerSerProGluSerProLysProAlaProAla 446
282 ACCGGAAGACATAGAACAATGTTCAAGCGGTACACGCGTGGGAACATG 331
447 ProGlySerHisArgGluLeuSerSerSerProThnSerLysAsn.... 461
332 TGCAGAGGCTTTGGACAACACAGAAAGGCTCTATTTCAT...CACGCCG 378
462 ....ArgSerHisGlyArgAlaLysArgAspLysSerHisSerHisThrP 477
379 CACATCGCAGCTACGATTGGCGGAGCTATCATCAGCCAGCAGCTCC 428
477 rSerArg.....ArgMetGlyArgSer...ArgSerProAlaThrAla 490
429 .....GTTCGCGCTGACCGC...CATGTACA 451
491 LysArgGlyArgSerArgSerArgThrProThrLysArgGlyHisSerAr 507
452 AACCGCGCAAAATCAAGCATAGCAAAATCAATGCA..... 488
507 gSerArgSerProGluThrPArgArgSerArgSerAlaGlnArgTrpGlyA 524
488 ..... 488
524 rSerArgSerProGluThrArgGlyArgSerArgSerProGluThrArgPro 540
489 GCGCGGAGGTTGCGG.....CAAGGAACAAAGCGGCTTACCAACA 532
541 GlyTrpSerArgSerArgAsnThrGlnArgGlyArgSerArgSerAl 557
533 TACAAGGGGTCAACAAATCATCAAGCGCTGCTTC...GGCGGAAGA 579
557 aArgArgGlyArgSerHisSerArgSerProAlaThrArgGlyArgSerA 574
580 ACATTCGT...CTGCGCGCA.....CCAGTCCCTCCCTCA 614
574 rSerArgThrProAlaThrArgGlyArgSerArgSerArgThrProAla 590
615 AGAAGCGGAGGAGCGCTATGGGTGATTTCTTGCAAAACGTCGCATA 664
591 ArgArgArgSerArg..... 595
665 CCAATGACGCTGCGGCAAAATTTGGCACACGTCAAAAGCGTGAAACCGCTG 714
596 .....SerArgThrProThrArgArgArgSerArgSer.... 606
715 TTTTTCGTGCGAAGCGCTGCGGAGCAAGGTTTCATTTGCAAT 764
607 .....ArgThrProAlaThrArgGlyArgSerArgSerArgThr 619
765 CCGCGCCCTCAAGGGAATTAACGCGCAAAAGCCATGATGCGCGC 814

```

```

620 ProAlaThrArgArg.....SerArgThrArgSerProValArgArgAr 634
815 T 815
634 g 634

```

seq_name: sp_human:Q9UQ35

seq_documentation_block:

AC Q9UQ35; PRELIMINARY; PRT; 2752 AA.

DT 01-MAY-2000 (TrEMBLrel. 13, Created)

DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)

DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)

DE RNA BINDING PROTEIN.

OS Homo sapiens (human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

NCBI_Taxid=9606;

RP SEQUENCE FROM N.A.

RA Ohtaki S., Umeki K., Sawada Y.;

RT "Homo sapiens mRNA for RNA binding protein, complete cds.";

RL Submitted (JUL-1998) to the EMBL/Genbank/DBJ databases.

DR EMBL; AB016092; BAA83718.1; -

DR InterPro: IPR002965; P. rich. extensn.

DR PRINTS: PR01217; PRICHEXTENS.

SO SEQUENCE 2752 AA; 299672 MW; 109C64F181097123 CRC64;

alignment_scores:

Quality: 128.50 Length: 317

Ratio: 0.905 Gaps: 18

Percent Similarity: 44.795 Percent Identity: 26.814

alignment_block:

US-09-303-518D-569 x Q9UQ35 ..

Align seg 1/1 to: Q9UQ35 from: 1 to: 2752

```

27 TCCGCCCTTTGCGAAGCGCCATGCATCCGTGTGACCGCCCTGCTCAAT 76
351 SerProSerProGluArgSerSerThnGlyProGluProProAlaProTh 367
77 GCGCTCCCTGCTGCGCGCTTCTGCTGTCACACAGCGTGGGAACCG... 122
367 rProLeuLeuAla.....GluArgHisGlyGlySerProGluP 380
123 ....GCTCGGACA.....TCTGGC 137
380 rLeuAlaThrThrProLeuSerGlnGluProValAsnProProSerGlu 396
138 GTTTTACCTTTAAAGAGA.....CCGCGCGCCATCGTCCCAATA 181
397 AlSerProThrArgAspArgSerProProLysSerProGluLysLeuPr 413
182 TGGGTAGCGAGCATGAATCCCGACCCCAAAACGGTCAAGCGCTTTT 231
413 ogInSerSerSerGluSerSerProProSerProGluProThLysV 430
232 GCGGAACGCAAAAGCGGTTTGGAACTTCCCGCGCTTTTTCAGAA 281
430 alSerArgHisAlaSerSerSerProGluSerProLysProAlaProAla 446
282 ACCGGAAGACATAGAACAATGTTCAAGCGGTACACGCGTGGGAACATG 331
447 ProGlySerHisArgGluLeuSerSerSerProThnSerLysAsn.... 461
332 TGCAGAGGCTTTGGACAACACAGAAAGGCTCTATTTCAT...CACGCCG 378
462 ....ArgSerHisGlyArgAlaLysArgAspLysSerHisSerHisThrP 477
379 CACATCGCAGCTACGATTGGCGGAGCTATCATCAGCCAGCAGCTTCC 428

```

```

477 roserArg.....ArgmetGlyArgSer...ArgSerProAlaThrAla 490
429 .....GTTCCCGCTGACCGC...CATGTACA 451
491 LysArgGlyArgSerArgSerArgThrProThrLysArgGlyHisSerArg 507
452 AACCCCGCAAAATCAAGCATAGACAAATCATGCA..... 488
507 gSerArgSerProGlnTrpArgArgSerArgSerAlaGlnArgTrpGlyA 524
488 ..... 488
524 rSerArgSerProGlnArgArgGlyArgSerArgSerProGlnArgPro 540
489 GGGGGGACGGTTCGGG.....CAAGGAAACCGCGCCCTACACAGA 532
541 GLYTrpSerArgSerArgSerArgThrGlnArgArgGlyArgSerArgSerAl 557
533 TACAGGGGTCAAAACAATCATCAAGCCCTGCGTTC...GGCGGAAGCA 579
557 aArgArgGlyArgSerHisSerArgSerProAlaThrArgGlyArgSerA 574
580 ACCATGCT...CCTGCCCGA.....CCAGTCCCTCCCTCA 614
574 rSerArgThrProAlaArgArgGlyArgSerArgSerArgThrProAla 590
615 AGAGCGGGGAGCGGTATGGTGATTTCTTCGGCAACCTGCCTATA 664
591 ArgArgArgSerArg..... 599
665 CCATAGCGCTGGCGCAAAATTTGGCACACGTCAAGCGCTGAACACCTG 714
596 .....SerArgThrProThrArgArgArgSerArgSer.... 606
715 TTTTCTGCTGCAAGCGCTGCTGCGAGACAGATTTCGATTGGACAT 764
607 .....ArgThrProAlaArgArgGlyArgSerArgSerArgThr 619
765 CCGCCCGCTCCAGAGGGAATTAACGGCGACAAAGCCCATGATGCCCGC 814
620 ProAlaArgArgArg.....SerArgThrArgSerProValArgArgArg 634
815 T 815
634 G 634
seq_name: sp_plant:Q9ARV4
seq_documentation block:
ID Q9ARV4 PRELIMINARY; PRT; 395 AA.
AC Q9ARV4;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DE 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE P0503E05.24 PROTEIN.
GN P0503E05.24.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBL_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. NIPPONBARE;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC
RT clone:P0503E05."
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP003021; BAB40020.1;
SQ SEQUENCE 395 AA; 42387 MW; 57DBE7AA401412C CRC64;

```

alignment_scores: 128.00 Length: 316
 Quality: 0.948 Gaps: 15
 Percent Similarity: 42.722 Percent Identity: 26.582

alignment_block:
 US-09-303-518D-569 x Q9ARV4 ..

Align seg 1/1 to: Q9ARV4 from: 1 to: 395

```

36 GCGAAGCGCCATGCACATCTGTTGACCGCCCTCTCAATTCGCTCC 85
  :::::::::::::::::::::
65 SerAsnArgHisLysIleGlnArgSerSerProThrProThrProLeu 81
86 TGCTGCCCTTTCCTGCTGTCACACGCTGGGAAACCGGCTCGACATCTG 135
81 s..... 81
136 GCGTTTACCTTTTAAAGAACCGCGCGCATCTCGTCCGAATATGCG 185
  :::::::::::::::::::::
82 .....ProThrAspHisPheProArgArgHisArgGlnProAspPhe 95
186 TCAGCAGCATGAATCCCGACCCCAAGGTCACAAACCGCTTTTTCGG 235
  ::::: ::::: ::::: :::::
96 AlaGlyLeuArgAlaArgArgGlyGlnArgGly.....CysGly 108
236 AAACGCG.....AAAGCGGTTTGA.....ACTGCCCC 267
  ::::: ::::: ::::: :::::
108 yAspGlyAspAspArgArgArgGlyGlyGlyGlyGlyArgSerCysGly 125
268 GCGTTTTCAGAAACCGGAGACATAGAACATGTTCAAAAGCGGTACA 317
  ::::: ::::: ::::: :::::
125 LnaIAlaGlnGlnGlyArgArgGlyArgSerSerValArgAlaProPro 141
318 CGCGTGGGACA.....TGTCAGCAGCGCTTGGCAACA 352
  ::::: ::::: ::::: :::::
142 LeuAlaIleThrHisLysAlaHisGlnAlaIleAlaGlyGlyAlaThr 158
353 A.....CGAAGGCTGCTAT..... 367
  :::::::::::::::::::::
158 rGlyArgArgGlnArgAlaIleAlaGlyProGlyGlyGlyArgSerA 175
368 .....TCATCAGCGCCGACA.....TCGCACTACGAT 397
  :::::::::::::::::::::
175 spGlyArgCysThrArgArgThrValLeuProGlnAspArgAlaThrThr 191
398 TGGCGGACGCTACATCAGCAGACGCTTCGTCGCCGTACGCCCATG 447
  ::::: ::::: ::::: :::::
192 TrpGly...AlaIleArgGlnSerArgArgGlyAlaGlyLysHisV 207
448 TACAAACCGCCGAATCAAGCGATAGACAAATCATCGACGGGCGAG 497
  ::::: ::::: ::::: :::::
207 aIaIleGlnGlnGlnAspGlyAlaAspGlyGlnArgAspGlyGlyGly 223
498 .....GGTTCGGGCAAGGAAACCGCGC 523
  :::::::::::::::::::::
224 AspAlaIleAlaValValValHisGlyArgArgSerArgAlaSerAla 240
524 CTACAGCATACAAAGGGTCAAAACAATCATCAACACCTCGCTGGGCG 573
  ::::: ::::: ::::: :::::
240 uAsnGlyValAlaGlyLysSerValThrAlaThrGlnAspAlaValGly 257
574 GAAGCAACCATCGTCCCTGCGCCGA..... 596
  :::::::::::::::::::::
257 rArgArgAlaGlyArgProValArgValAlaGlnGlyGlyValValThrGln 273
597 .....CCAGTCCCTCCCTCAAGAGCGGGGAGCGGTATGGTGAT 643
  ::::: ::::: ::::: :::::
274 LeuAspArgProSerProPheSerArgGlnGlyThrArgLysArgGly 290
644 TCTTCGCAACACGCTATACCATGACGCTGGCGCAAAATTTGGACAC 693
  ::::: ::::: ::::: :::::
290 yArgArgCysArgGlyArgGlyAspAspGlyGlyGlyGln.....A 304

```

```

694 GTCAGAGCGGTGAAACGCTGTTTCTGCTG.....CGA 728
||||| :||| :|||
304 rglntargProh1sArgArgSerTrileuLeuTrpTrglnAlaArg 320
729 ACAGCGTCCGTCGGGACAGATTTCATTCACAT.....764
||||| :||| :|||
321 gLyThrAlaAlaArgAlaProLeuLeuProAlaHisProThrGlnLeu 337
765 .CCGCCCCGTCAGAGGGAATTGAAGCGGACAAAGCCCATGATGC 809
||||| :||| :|||
337 uProProProMetProAlaAlaHisAlaGlnLeuAlaSerCys 352
seq_name: sp.bacteria:092182

```

```

seq_documentation_block:
ID 092182 PRELIMINARY; PRT; 323 AA.
AC 092182;
DT 01-MAY-1999 (TREMBLrel. 10, Created)
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
DT 01-MAY-1999 (TREMBLrel. 10, Last annotation update)
DE LIPID A ACTYLTRANSFERASE.
GN MSBB.
OS Salmonella typhimurium.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Salmonella.
OX NCBI_Taxid=602;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC14028;
RX MEDLINE=99116994; PubMed=9920266;
RA Low K.B., Ittensohn M., le T., Platt J., Sodi S., Amoss M., Ash O.,
RA Carmichael E., Chakraborty A., Fisher J., Lin S.L., Luo X., Miller S.,
RA Zheng L.-M., King I., Fawcett J., Bermudes D.;
RT "Lipid A mutant Salmonella with suppressed virulence and TNFalpha
RT induction retain tumor-targeting in vivo.";
RL Nat. Biotechnol. 17:37-41(1999).
DR EMBL; AF039020; AAD03801.1;
KW Transferase; Acyltransferase.
SQ SEQUENCE 323 AA; 37287 MW; 3E7B3F35E997C47 CRC64;

```

```

alignment_scores:
Quality: 127.50 Length: 290
Ratio: 0.885 Gaps: 10
Percent Similarity: 49.655 Percent Identity: 24.138

```

alignment_block:

US-09-303-518d-569 x 092182 ..

Align seg 1/1 to: 092182 from: 1 to: 323

```

64 GCCGTCGTCGAATGCTGCTCCCTGCGCGCTTCTGTCACAGCCT 113
||| :||| :|||
31 AlalalmetAlaGlylleAlaLeuThrProAlaSerPheArgAspPro 47
114 GGGAAACGGCTCGACATCTGGCTTTTACCTTTTAAGAGAGACCGCG 163
| :||| :|||
47 uLeuAlaThrLeuGlyArgPheAlaGlyArgLeuGlyLysSerSerAla 64
164 CCGCATGTCGCGCAATATGCT.....CAGGACGACATGAT 201
||| :||| :|||
64 rGArgAlaLeuLeuAlaSerLeuSerLeuSerPheProGlnArgSerGlnAla 80
202 CCGGACCCCAAAACGGTCGAAGCCGTTTTCGGAAGGACGAAAGGCGG 251
||| :||| :|||
81 GlnArgGlnAlaAlaIleValAspGlnMetPheAlaThrAlaProGlnAla 97
252 TTGGAACTGCCCCCGGCTTTTTCAGAAACCGGACACATAGAACAA 301
||| :||| :|||
97 tAlaMetAlaGlnAlaLeuAlaMetArgGlyProLysLysIleGlnGln 114
302 TGTTCAAAGCGGTACAGCGCTGGGACACATGTCAGACGCGCTTGGACAA 351

```

```

114 rGValAsp...TrpGlnGlyLeuGlnIleIleGlnIleMetArgArgAsn 129
||| :||| :|||
352 CAGGAGGCGTGTATTCATCAGCGGCGCATCGGACGCTAGCATTTGG 401
||| :||| :|||
130 AspGlnLysValIlePheLeuValProHisGlyTrpGlyValAspIle 146
402 CCGAGCGTATCAGCAGCAGCAGCTTCGTCGCGTACCGGACGATACA 451
||| :||| :|||
146 oAlaMetLeuMetAlaSerGln...GlyGlnLysMetAlaAlaMetPhe 162
452 AACCGCGGAAATCAAGGATAGCAAAATCATGACAGCGGCGGCGGTT 501
||| :||| :|||
162 tAsnGlnGlnLysProValPheAspTrpTyrIleTrpAsnThrValArg 178
502 CCGGCGCAAGGAAACCGCGCTACAGCATACAGGCGGTCAAAATAT 551
||| :||| :|||
179 ArgPheGlnGlyArgLeuHisAlaArgAsn...AspGlyIleLysPro 194
552 CATCAAGCGCGTGTGCGGCGGAAAGCAACATCGTCTGCGC.....G 595
||| :||| :|||
194 eIleGlnSerValArgGlnGlyTyrTrpGlyTyrTyrLeuProAspGln 211
596 ACCAGTCCCGTCCCGTCAAGAGCGGCGGAAAGCGGTATGGGTGATTC 645
||| :||| :|||
211 sPheIleGlyProLysIleSerGlu.....PheValAspPhe 222
646 TTGGCAAAACGCTGCTATACATGACGCTGGCGCAAAATTTGGACAGT 695
||| :||| :|||
223 PheAlaThrTyrLysAlaThrLeuProAlaIleGlnArgLeuMetLys 239
696 CAAGCGCGTGAACCGCTGTTTCTGCTGCGAAGCGCTG.....735
||| :||| :|||
239 L.....CysArgAlaArgValIleProLeuP 248
736 .CCTGGCGGACAGGT.....TTGCATTTGCATTCGCGGCC 771
||| :||| :|||
248 heProValTyrAsnGlyLysThrHisArgLeuThrIleGlnIleArgPro 264
772 GTCGAGGCGGAA...TGAACGGGACAAAGCCCATGATCGCGCGTGT 818
||| :||| :|||
265 ProMetAspAspLeuLeuThrAlaAspAspHisThrIleAlaArgAr 281
819 CAACCGCAATGCGCAATATGATGATGCGCGTTCGACAGCAGATGCT 868
||| :||| :|||
281 tAsnGlnGlnValGlnIlePheValGlyProHisProGlnGlnTyrThr 298
869 TTATGTACAAACCGCTACAA 888
||| :||| :|||
298 rPheLeuLysLeuLeuLys 304

```

```

seq_name: sp.human:016824
seq_documentation_block:
ID 016824 PRELIMINARY; PRT; 797 AA.
AC 016824;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE PROFLAGGRIN (FRAGMENT).
GN FLAG.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91064347; PubMed=2248957;
RA Gan S.O., McBride O.W., Idler W.W., Markova N., Steinert P.M.;
RT "Organization, structure, and polymorphisms of the human profilaggrin
RT gene [published erratum appears in Biochemistry 1991 Jun
RT 11:30(23):5814].";
RL Biochemistry 29:9432-9440(1990).

```

DR EMBL: M60502; AAA63248.1; -
 DR InterPro: IPR003303; Filaggrin.
 DR PRINTS: PR00487; FILAGGRIN.
 FT NON_TER
 SQ SEQUENCE 797 AA: 8516 MW: 60E6184763BDA86B CRC64;

alignment_scores:
 Quality: 125.50 Length: 222
 Ratio: 1.037 Gaps: 10
 Percent Similarity: 54.505 Percent Identity: 24.324

alignment_block:
 US-09-303-518D-569 x Q16824 ..

Align seg 1/1 to: Q16824 from: 1 to: 797

```

162 CGCGCGCATGCTGGCAATATGCTGACGAGGAGCATGATCCGACCCCA 211
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
266 ArgThrSerAlaArgGlnGlySerSerValSerGlnSprSerI 282
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
212 AAAGGTCAAAAGCCGTTTTCGGAACGCGCAAAAGCGGTTTGAAGCTT 261
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
282 n...GlyHisSerGlu.....AspSerGluArgArgSerGlySerA 295
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
262 GCGCCCGCGGTTTTCAGAAAACCGAGACATAGAACAAATGTTCAAGC 311
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
295 LaserArg.....AsnHisArgGlySerAlaGlnGlu 305
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
312 GGTACACGCTGGGAACA.....TGTGCAGC 337
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
306 GlnSerAlaArgSerArgHisProArgSerHisGlnAspArgAl 322
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
338 AGGCTTTGA.....CAACACGAAAGGCTGCTGTTCATCAGC 375
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
322 aGlyHisGlyAspSerAlaGlnSerSerArgGlnSerGlyThrHisSA 339
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
376 CGGCACATCGGAGCTACGATTTGGCGGAGCTACATCAGCAGCAGCT 425
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
339 LagLysHisSerSerGlyGlnAlaAlaSerSerHisGlnAlaArg 355
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
426 TCCGTTCCCGCTGACCGCCATGTACAAACCGCCGAAATCAAGAGCATG 475
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
356 SerSerAlaGlnGlyAlaArgHisGlySerHisTyrGlnGlnSerAlaAspSe 372
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
476 ACAAAATATATGCGAGC.....GGCGAGGCTCGCGCAAAAGGAAGA 516
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
372 rSerArgHisSerGlyGlnGlyHisGlyGlnAlaSerSerAlaValArgA 389
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
517 ACCGCGCTACGACATACAGGGGT..... 542
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
389 spSerGlyHisArgGlySerSerGlySerGlnAlaSerAspAsnGlnGly 405
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
543 .....CAACAAATCATCAAGCCCGCTCGGGCGA.....AG 577
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
406 HisSerGluAspSerAspThrGlnSerValSerAlaHisArgGlnAlaGln 422
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
578 CAACATCGCTCTGCCCGACAGCTCCCTCCCTCAAGAGCGGAGGAA 627
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
422 yArgHisHisGlnSerHisGlnGlnSerThrArgGlyArgSerArgGlyA 439
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
628 GG.....CGTATGGGTGATTTTTCGCAAACTGCTATACATGAC 671
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
439 rSerGlyArgSerGlySerPheLeuGlyGlnValSerThrHisGlnGln 455
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
672 GCT.....GGCGCAAAATTTGG 688
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
456 SerGlnSerAlaHisGlyArgAlaGlyProSerThrGlnGlyArgGlnGln 472
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
689 CACACGTCAAGGCGT 704
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
472 ySerArgHisGlnGln 477
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

seq_name: sp_bacteria:Q52747

seq_documentation_block:
 ID Q52747 PRELIMINARY; PRT; 680 AA.
 AC Q52747;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE ENDO-GLUCANASE.
 OS Ruminococcus flavefaciens.
 OC Bacteria; Firmicutes; Bacilli/Clostridium group; Clostridiaceae;
 OC Ruminococcus.
 OC NCB1_TaxID=1265;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=186;
 RA Huang C.M., Asmundson R.V., Yu P.L.;
 RT "Nucleotide sequence of a cellulase gene complex from Ruminococcus
 flavefaciens strain 186 coding for multi-cellulase activities.";
 RL Submitted (SEP-1990) to the EMBL/GenBank/DBJ databases.
 DR EMBL: X56082; CAA39559.1; -
 SQ SEQUENCE 680 AA: 75503 MW: 5A589F334846D6F CRC64;

alignment_scores:
 Quality: 125.00 Length: 317
 Ratio: 0.919 Gaps: 21
 Percent Similarity: 42.902 Percent Identity: 28.076

alignment_block:
 US-09-303-518D-569 x Q52747 ..

Align seg 1/1 to: Q52747 from: 1 to: 680

```

29 CCCCTTGGCAACGCCATGACATCTGTTGACCG...CCCTGCTCAAA 75
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
27 ProThrProProSerProSerThrGlnSerSerProThrProSerSerPr 43
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
76 TGCTCTCC.....TGTGCGCG.....TTTCG 101
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
43 oAlaSerProArgSerProThrValArgGlyArgGlyAlaValAlaProA 60
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
102 TGTGACACGCTGGGAACCGCGCTCGACATCGCGTTTACCTTTAA 151
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
60 Lcys.HisAla.....AspAlaArgThrAlaGly.....Ar 70
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
152 AGGAGA...CGCGCGCGCATGTCGCCAATATGCGTACGAGGAGCAG 198
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
70 gGlyArgAspProArgHisGlnGlnCysAspAlaArgLeuArgGlyArgHis. 86
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
199 AATCCGACCCCAAAACGTCMAAGCGGTTTTTGGGAAACGGCAAAAG 248
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
87 .....ArgProArgLeuGlyHisAlaValAlaValArgAspArgGly..... 99
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
249 CGGTTTGAACTTGCCCGCGCTTTTTCAGAAAACCGAGACATAGAAA 298
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
100 .....LeuGlyProHisGlnAr 105
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
299 CAATGTTCAAGCGGTACAGCGCTGGAAATGTGCA.....GCAAGGT 342
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
105 gAspCysValArgGlyAlaCysHisGlyProSerValHisArgAlaAspT 122
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
343 TTGGCAAAACGAAAGGCTGTATTCATCAGCCGCGACATCGGAGCTA 392
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
122 rHisGlnProArgArgAlaAlaLeuGlu..... 131
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
393 CGATTTGGCGGAGCTACATCAGCCAGCAGCTTCCTCCGCTGACCG 442
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
132 LeuValGlyGlnLeuValHisProGlnSerGlyLeuValProValAlaAr 148
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
443 CCATGTA.....CAACGCGCGAAATCAAGC.... 470
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
148 gGlyValProAlaProProProArgAspArgArgArgAspGlnSerArgP 165
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```



```

471 .....GATAGCAAAAT... 482
165 roATgATgAlaLeuValGlnLeuProLeuProArGAspArgGlnGlnPro 161
483 .....CATGCAAGCGGCAAGG 499
182 GLyGlnHISGlnProHISAspAlaAspAlaValHISGlyGlnGlyArgGln 198
500 TTGCGG...CAAGGAAAAACCGCGCTACAGCATACAGAGGCTCAAA 546
198 YAlArGGLyGlnLeuGlnSerArgArgProArGAspArgAlaValArG 215
547 CAATATCATCAAGCCCTGCGTGGCGGCAAGCAACCATGCTGCTGCCCA 596
215 sPHISArGAspGLyProSerAspGLyArGSerPheAspProArGAspArg 231
597 CCAAGTCCCTCCCTCAAGAGCGGCGGCAAGCGCTATGGGTGGATTCT 646
232 GlnArGAlaVal...ArGArgGlnArGGLyArGArg..... 242
647 TCGGCAAACTGCTATACATGACGCTGCGGCAAAATTGGCAACAGTC 696
243 .....HisArGLeu.....ValAspArgArgV 250
697 AAGG...CGTGAACCCCTGTTTCTGCTGCGAAGCGCTGCGGCGG 743
250 aArGProArGArgGlnProValProArGValArGArgGlnArgGly 266
744 ACNAGGTTCGATTTGCATCGCGCGCGCTGCAAGGGA..... 782
267 ValArG.....AlaHISGlnHISArGArgArgArgGlyProLeuAl 280
783 .....ATTGAAGCGCGCAAGAACCCCATGATGCGCGCTGTCA 821
280 aGlyAspLeuValGlnArGAlaGlnAlaHISProArGSerValGln 296
seq_name: sp_bacteria:Q53478

```

```

seq_documentation_block:
ID Q53478 PRELIMINARY; PRT; 456 AA.
AC Q53478;
DT 01-NOV-1996 (TREMBLrel_01, Created)
DT 01-NOV-1996 (TREMBLrel_01, Last sequence update)
DT 01-DEC-2001 (TREMBLrel_19, Last annotation update)
DE ORF2 PROTEIN.
GN ORF2.
OS Neisseria gonorrhoeae.
OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
OX NCBI_TaxID=485;
RN [1]
RP SEQUENCE FROM N. A.
RX MEDLINE=95020543; PubMed=7934834;
RA Facius D., Meyer T.F.;
RT "A novel determinant (coma) essential for natural transformation
RT competence in Neisseria gonorrhoeae and the effect of a coma defect on
RT plin variation."
RL Mol. Microbiol. 10:699-712(1993).
DR EMBL; S75490; AAB32262.1;
SQ SEQUENCE 456 AA; 54380 MW; 30AC7B5CDFB912D5 CRC64;

```

```

alignment_scores:
  Quality: 124.00      Length: 343
  Ratio: 0.800        Gaps: 26
  Percent Similarity: 45.190  Percent Identity: 25.656

```

```

alignment_block:
US-09-303-518D-569 x Q53478

```

```

Align seg 1/1 to: Q53478 from: 1 to: 456

```

```

32 CTTTGGCAACCGCATGACATCTCTTGAACCGCGCTGCTCAANTGCGCTC 81

```

```

143 ValGysArGPro.HISLysHISArGAlaLeuPro...ArgLLeProA 158
82 TCCSTGCTGCGCGCTTCTGCTGCACAGCGTGGGAACCGCGCTGGACA 131
158 snArGAsnArGSerPheValArGArgLysCys.....PheArGThr 171
132 TCTGGCGTTTACCTTTT.....AAGGAAGACCGCGC..... 164
172 ValGlyTyrLysProCysGlnGlnGlnHISArGProArGThrGlnThrGln 188
165 .....GCGCATGCTGCGCAATATGCGTACAGGACGATGATATCCGACCC 210
188 nAlaAlaArGInGlnGlnGlnGlnSerGlyThrHISArGInHISGln 205
211 AAAGCGTCAAGCCGTTTTCGCGAAGCGGCAAA.....AGCGGTGT 254
205 In.....ArgGlnArGArgGlyGlnGlnGlyLysLeuArGArgHISLys 219
255 GGA.....ACTTGGCCCG 268
220 GLyLysProHISGlnLysProGlnSerValPheArGlnLysCysProAr 236
269 CGTTTTCGAAGAACCGGAGACATAGAACATGTTCAAGCGGTACAC 318
236 gLeuLeuLysArGArgLysArGHis..AspArgThrGlnGlnArGAla 252
319 GCGTGGACAT.....GTGACAGAGC 341
252 Gg.GlyGlnHISProArGArgLysProArGArgLysArGHisArGAla 268
342 TTTGACAAACACGAGGCGTGTATTCATCACCGCGCACATCGGCAAGT 391
268 gAlaAspGlnArGln..... 273
392 ACGATTGGCGGCAAGCGTACATCACCGACAG..... 423
274 .....ArgGlyLysGlnLLeThrGlnGlnGlnGlnArGInHISArG 287
424 .....CTTCCGTTCCCGCTG.....AC 440
288 SerArGProLeuProAlaHISGlyGlyLeuProPheProLeuValGlnThr 304
441 CGCATGTACA.....AACCGGAAATCAAGCGATGACGACAAA 461
304 r.ProCysThrArGProTyrGlnGlnLys..HISGlnThrGlnArGlnPr 320
482 TCATGACAGCGGCGAGGTTCCGCGCAAGAAAAAACCGCGCTACAGC 531
320 oSerThrAspThrGln...AsnArGlnArGValGlnGln.....Gln 334
532 ATACAGGCGTCAACAAATCATCAAGCCCTGCGTGGCGGCAAGCAAC 581
334 IsArGArgGlnArGLeuProArGArgPro..... 344
582 CATGCTGCTGCGGACCGACGCTCCCTCCCTCAAGAGCGGGAAGGCG 631
345 ArgArGlnSerSerProAlaProArGProArGInArGArgHISGln 361
632 TATGGGTGATTTCTTCCGCAAACTGCTATACCATGACGCTGCGGCA 681
361 nAsnThrAlaLeuArGArg.....HISGlyLysThrGlyL 373
682 AAATGGCAACGTCAA.....AGCGTGAACACCTGTT 716
373 YsGlnArGValLysGlnArGAlaArgThrProArGArgGlnAsnPro... 388
717 TTTCTGCTGGAAGCGCTGCTGCGGACAGGTTTGCATTGGACATGCG 766
389 .....ArgAlaArGProGlyGlyLysAlaArGGLyThrGlyLysGlnPr 403
767 G.....CCCGTCA.....AGGGAATTGAACGCG 792

```

403 OLeuArgGlnProArgLySGlnHisArgHisHisArgAsnValGlnThr 420
 793 GACAAGACCCATGATGCCCGCTGTCA..... 821.
 420 rGAsnThrAspGlnMetArgGlnProGlnArgProGlnArgLeuProGln 436
 822CCGCAATGCCGAATA 836
 437 SerGlyAlaGlnCysArgLeu 443

seq_name: sp_bacteria:09LCV5

seq_documentation_block:
 ID 09LCV5 PRELIMINARY; PRT; 886 AA.
 AC 09LCV5;
 DT 01-OCT-2000 (Tremblurel. 15, Created)
 DT 01-OCT-2000 (Tremblurel. 15, Last sequence update)
 DT 01-JUN-2001 (Tremblurel. 17, Last annotation update)
 DE ATP OPERON (FRAGMENT).
 OS Rhodospirillum rubrum.
 OC Bacteria; Proteobacteria; alpha subdivision; Rhodospirillaceae;
 OC Rhodospirillum.
 OX NCBI_TaxID=1085;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=65251588; PubMed=2861810;
 RA Falk G., Hampe A., Walker J.E.,
 RT "Nucleotide sequence of the Rhodospirillum rubrum atp operon."
 RL Biochem. J. 228:391-407(1985).
 DR EMBL; X02499; CAB97255.1; .
 DR InterPro: IPR001899; Gram_pos_anchor.
 DR PostPro: PS00343; GRAM_POS_ANCHORING; UNKNOWN_1.
 FT NON_TPR 886
 SQ SEQUENCE 886 AA; 95954 MW; BAEL40D8802E6FC CRC64;

alignment_scores:
 Quality: 123.50 Length: 433
 Ratio: 0.787 Gaps: 21
 Percent Similarity: 36.259 Percent Identity: 21.478

alignment_block:

US-09-303-518d-569 x 09LCV5

Align seg 1/1 to: 09LCV5 from: 1 to: 886

51 CATCCGTTGACCG..... 65
 21 HisProThrGlyArgThrGlyLeuGlyAspArgLeuLeuIleAlaIle 37
 66CCTGCTCAATGCCCTCTCCCTGCGCTTCTCTGTCGACACG 111
 37 tValThrAlaGlnAlaProIleAlaValAlaPheAsnGlnProSerGlyA 54
 112 CTGGGAACCGGCTCGACATCTGGCTTTTACCTTTAAAGAGACCG 161
 54 tAlaIleArgAlaPheGlnProMetThrAlaAlaProThrGlnGlyGlnArg 70
 162 CGC.....GCCATCGTCG 175
 71 ArgIleAlaProAlaIleGlyIuysGlnGlnAsnLeuLeuAlaArgArg 87
 176 CCAATATGCTCAGCGACGATGATCCGACCCCAAAACGTCGAAGC 225
 87 nGlyPheLeuAspGlyGlyAspGlnArgArgAlaGlnProAlaGlnAlaG 104
 226 GTTTTGGGAAACGGCAAAAGGCTTGGAACTGCGCCCGCTTTT 275
 104 tArgArgAlaGlyHisGlyArgArgArgGlyArgArg..... 115
 276 CAGAAACCGGACGACATAGAAACAAT..... 302
 116 ArgArgGlyGlyArgArgArgArgGlyGlyArgPheLeuLeuGlyLeu 132

303GTTCAAGCGGTACACGCGTGGGACACATGTGCACAG 339
 132 uValValArgProAlaGlyArgGlyGlnThrLeu.....ProAlaG 147
 340 GCTTGGACAAACACGA..... 356
 147 tSerGlyLysIleArgGlyGlnProPheLeuArgArgGlyAlaAlaGln 163
 357AGGCTGTATTCATCACGCGC 379
 164 lGluGlnGlyAspLeuGlyHisPheArgSerAlaValAlaThrGlyG 180
 380 ACATCG.....CAGC 390
 180 HisGlnMetProIleAlaAlaLeuAlaGlyAlaLeuAspGlyPheGlnA 197
 391 TAGCATTTGGCGGACGCTACAT...CAGCCAGCAGTTCGCTCCGCT 437
 197 rGArgArgGlyArgGlyGlnHisGlyGlyLysAlaAlaGlnAlaGlyAla 213
 438 GACCGC...CATGTACAAACCGCGCAAAATCAAGC..... 470
 214 AspHisGlyHisValAlaGlyValIleArgHisAlaLeuLeuPheG 230
 471GATGACAAATCATCGACGCGGACG 498
 230 uGlyAlaValAlaPhePheIleAspAspArgGlnAlaGlnIleGlyIleG 247
 499 GTTCGGCGCAAGAAACCGCGCTACAGATCAAGGAGGTCAACA 548
 247 tGlnIysGlnIleArgArgSerArgProHisHisArgArg..... 260
 549 AATCATCAAGCCCTGCTGGCGGACCAACCATGCTCTCCGACG 598
 261ProGlnGly.....HisArgProSerPr 270
 599 ACGTCC.....CTCCCTCAGAGCGGCGA. 626
 270 oProProArgArgArgAlaHisLeuArgMetProLeuHisArgSerGlyA 287
 627AGCGTATGGGTGA..... 641
 287 lGluAlaValGlyGluAlaProArgProLeuGlyGlnGlyAspPhe 303
 641 641
 304 GlyGlnGlnHisGlnSerLeuProProAlaAlaGlnGlyPheGlyAsp 320
 642TTTCTTGGC 651
 320 gLeuGlnIleAspLeuGlyLeuAlaArgAlaGlyAspProPheGlnGln 337
 652 AAACCTGCTATACCATGACGCTGGCGGCA..... 663
 337 lInThr.....GlyLysSerLeuValLysArgArg 346
 684ATTGGCACAGTCGAAGCGCTGA 706
 347 AlaglnAlaValArgArgArgLeuLeuIleGly...ArgGlnGlnArg 362
 707 AAACCTGTTTTTCTG.....CTGCAGACGCTGCGCGGACAAAGT 750
 362 rGlyProValGlyIleGlyGlyLeuArgArgProPheIleArgThrGly 379
 751 TTGATTTGGACATCG.....CCCGGCA 776
 379 tArgLeuGlnGlnProGlyLeuGlyHisArgArgHisHisArgArgPro 395
 777 AGGGA..... 782
 396 HisGlyArgPheProArgGlnIleGlyAlaGlyGlnGlyGlnThrGly 412

